NECA 2018

신의료기술평가 권고안 재정립 연구

NECA 한국보건의료연구원
National Evidence-based Healthcare Collaborating Agency

NECA - NP 18-009

신의료기술평가 권고안 재정립 연구

2019. 5. 31.

NECA 한국보건의료연구원
National Evidence-based Healthcare Collaborating Agency

주 의

1. 이 연구는 한국보건의료연구원 연구윤리심의위원회 승인(NECA IRB 19-001)을 받은 연구사업입니다.

2. 이 보고서는 2018년도 정부(보건복지부)의 재원으로 한국보건의료연구원에서 수행한 연구사업(과제번호: NP18-009)의 결과보고서로 한국보건의료연구원 연구기획관리위원회(또는 연구심의위원회)의 심의를 받았습니다.

3. 이 보고서 내용을 신문, 방송, 참고문헌, 세미나 등에 인용할 때에는 반드시 한국보건의료연구원에서 시행한 연구사업의 결과임을 밝혀야 하며, 연구내용 중 문의사항이 있을 경우에는 연구책임자 또는 주관부서에 문의하여 주시기 바랍니다.

연 구 진

연구책임자

이유경 순천향대학교 부천병원 진단검사의학과 교수

고려진 한국보건의료연구원 신의료기술평가사업본부 부연구위원

참여연구원

신채민 한국보건의료연구원 신의료기술평가사업본부 선임연구위원

전미혜 한국보건의료연구원 신의료기술평가사업본부 부연구위원

신해원 한국보건의료연구원 신의료기술평가사업본부 부연구위원

김주연 한국보건의료연구원 신의료기술평가사업본부 부연구위원

김진형 한국보건의료연구원 신의료기술평가사업본부 주임연구원

박근영 한국보건의료연구원 신의료기술평가사업본부 주임연구원

손수경 한국보건의료연구원 신의료기술평가사업본부 주임연구원

차 례

요약문 ····· i

Ⅰ. 서론 ····· 1
1. 연구배경 및 필요성 ····· 1
2. 연구 목적 ····· 13

II. 연구 방법 ····· 14

III. 연구 결과 ····· 16
1. 주요국 및 기관의 근거수준·권고등급 체계 ····· 16
2. GRADE 체계 ····· 45
3. 신의료기술평가자료의 현황 분석 결과 ····· 72
4. 근거수준 및 권고체계 개선방안 ····· 94
5. 권고체계(안) 시범적용 ····· 102

IV. 결론 및 제언 ····· 116

V. 참고문헌 ····· 120

표 차례

표 1. 신의료기술평가결과 ····· 4
표 2. 질 평가 판정기준 ····· 4
표 3. 질 평가 결과 ····· 4
표 4. 근거의 수준 ····· 5
표 5. 조기기술 및 연구단계 의료기술 등급 부여 기준 ····· 5
표 6. 제한적 의료기술 분류 기준 ····· 6
표 7. 중재시술의 평가대상여부 심의 기준 ····· 6
표 8. 기타 검사의 평가대상여부 심의기준 ····· 7
표 9. 혁신의료기술 잠재성 평가 척도 ····· 8
표 10. SIGN에서 제시한 근거수준 체계 ····· 19
표 11. SIGN에서 제시한 권고등급 체계 ····· 20
표 12. NICE 의료기술평가 권고 ····· 21
표 13. 종양질환에서 PET 검사 사용 지침에 대한 근거수준 및 권고등급 ····· 24
표 14. 가이드라인 개발을 위한 근거의 수준 ····· 24
표 15. ACCP의 권고등급 체계의 기준 ····· 26
표 16. ACCP의 권고 등급 ····· 26
표 17. 권고강도에 미치는 영향 요인들 ····· 27
표 18. 이득, 위해, 부담, 비용 측정에서 고려되는 주요 요인의 유형 ····· 27
표 19. ACN의 근거수준 체계 ····· 33
표 20. NHFA/CSANZ의 권고등급 체계 ····· 33
표 21. 호주 NHMRC 근거수준 체계 ····· 34
표 22. MSAC에서 의료기술평가 결과에 대한 권고안 ····· 35
표 23. HTA 프로그램 시행 절차 ····· 37
표 24. USPSTF의 근거수준 고려요소 ····· 38
표 25. USPSTF의 권고등급 체계 ····· 39
표 26. 미국의 급여결정 결과 ····· 40
표 27. 주요국 의료기술평가 결과 비교 ····· 41
표 28. 주요국의 의료기술평가 권고결정 고려요인 ····· 42
표 29. 근거의 수준 유형 ····· 44
표 30. 권고의 수준 유형 ····· 44
표 31. GRADE에서 근거수준 의미 ····· 47
표 32. GRADE에서 근거수준 평가기준 ····· 48

표 33. GRADE의 권고 강도 4가지 ····· 49
표 34. 권고등급을 결정하는 요인 ····· 50
표 35. 의사결정 유형별 EtD frameworks 기준 ····· 50
표 36. Evidence to Decision(EtD) frameworks 기준에 대한 질문 및 세부지침 ··· 53
표 37. Coverage를 위한 EtD framework (질문섹션) ····· 58
표 38. 근거에 기반한 의사결정 틀(Evidence to Decision Frameworks) ····· 59
표 39. 평가결과 및 기술유형별 신의료기술평가 전체 신청 현황 ····· 73
표 40. 기술유형별 신의료기술 평가 결과 분포 ····· 73
표 41. 기술유형별 평가방법에 따른 신의료기술 평가 결과 ····· 75
표 42. 평가결과별 근거의 수준 유형 ····· 77
표 43. 처치 및 시술에서 평가결과별 근거의 수준 ····· 78
표 44. 체외진단검사에서 평가결과별 근거의 수준 ····· 78
표 45. 영상검사에서 평가결과별 근거의 수준 ····· 79
표 46. 기타검사에서 평가결과별 근거의 수준 ····· 80
표 47. 신의료기술로 인정된 기술의 급여 유형별 근거의 수준 ····· 81
표 48. 신의료기술로 인정된 기술 중 급여 결정 판단 근거가 공개된 기술의 급여 현황 ····· 82
표 49. (처치 및 시술)급여유형별 급여 결정 시 의결사항 특징 ····· 85
표 50. (체외진단검사)급여유형별 급여 결정 시 의결사항 특징 ····· 86
표 51. (영상검사)급여유형별 급여 결정 시 의결사항 특징 ····· 87
표 52. (기타검사)급여유형별 급여 결정 시 의결사항 특징 ····· 87
표 53. 근거의 수준이 A이면서 비급여로 결정된 기술의 사례분석 ····· 90
표 54. 근거의 수준이 D이면서 급여로 결정된 기술의 사례분석 ····· 91
표 55. 근거 수준 평가를 수행하지 않은 기술이면서 급여로 결정된 기술의 사례분석 · 92
표 56. 근거 수준 평가를 수행하지 않은 기술이면서 급여로 결정된 기술의 사례분석 ·· 93
표 57. 권고체계(안) ····· 96
표 58. 신의료기술평가 권고등급 평가기준(1안) ····· 98
표 59. 신의료기술평가 권고등급 평가기준(2안) ····· 99
표 60. 평가기준(안) 비교 ····· 100
표 61. 신의료기술평가 내용(항PLA2R IgG 항체검사 [ELISA][IIFA]) ····· 103
표 62. 권고등급 논의결과(항PLA2R IgG 항체검사, 평가항목 1안 적용) ····· 105
표 63. 권고등급 논의결과(항PLA2R IgG 항체검사, 평가항목 1안 적용) ····· 107
표 64. 신의료기술평가 내용(툴륨레이저를 이용한 전립선 광적출술) ····· 109
표 65. 권고등급 논의결과(툴륨레이저를 이용한 전립선광적출술, 평가항목 1안 적용) 113
표 66. 권고등급 논의결과(툴륨레이저를 이용한 전립선광적출술, 평가항목 2안 적용) 115

그림 차례

그림 1. 신의료기술평가 절차 ····· 2
그림 2. 혁신의료기술 별도 평가 절차 ····· 8
그림 3. 체외진단검사 등재 절차 개선 ····· 10
그림 4. NECA의 의료기술재평가 기본체계도(안) ····· 11
그림 5. The National Institute for Health Research HTA program ····· 18
그림 6. 호주 시장 진입 및 급여 결정 시 의료기술평가 절차 ····· 31
그림 7. 권고안 도출 관련 GRADE 과정의 도식화 ····· 46
그림 8. 의료중재에 대한 통합된 의사결정 framework ····· 57
그림 9. 평가결과별 기술유형 분포 ····· 73
그림 10. 기술유형별 신의료기술 평가 결과 분포 ····· 74
그림 11. 기술유형별 평가방법에 따른 신의료기술 평가결과 분포 ····· 76
그림 12. 평가결과별 근거의 수준 유형 ····· 77
그림 13. 처치 및 시술에서 평가결과별 근거의 수준 ····· 78
그림 14. 체외진단검사에서 평가결과별 근거의 수준 ····· 79
그림 15. 영상검사에서 평가결과별 근거의 수준 ····· 79
그림 16. 영상검사에서 평가결과별 근거의 수준 ····· 80
그림 17. 신의료기술로 인정된 기술의 급여 유형별 근거의 수준 ····· 81
그림 18. 급여기술에서 의사결정 판단 요소 비율 비교 ····· 83
그림 19. 선별급여기술에서 의사결정 판단 요소 비율 비교 ····· 84
그림 20. 비급여기술에서 의사결정 판단 요소 비율 비교 ····· 84
그림 21. 신의료기술평가 권고체계(안) ····· 95

요 약 문

▢ 연구 배경

신의료기술평가는 기본적으로 체계적 문헌고찰을 통해 문헌적 근거에 기반한 의학적 근거 수준을 평가하여 의료기술의 도입에 관한 의사결정과정을 수행하고 있고, 신의료기술평가 결과는 안전하고 유효한 '신의료기술'과 '연구단계기술'로 크게 구분되어 평가되고, 체계적문헌고찰을 수행한 경우에는 SIGN의 질평가 결과에 따른 문헌적 근거수준(A~D)이 함께 제시되고 있다. 신의료기술평가에서 신의료기술로 인정되어야 의료행위의 건강보험 급여 신청이 가능하고, 신의료기술의 평가결과는 건강보험과정안에서 급여기준마련의 근거자료로 활용되고 있다.

현재 신의료기술의 도입에 있어 신의료기술 평가단계와 건강보험 결정 단계로 이분화되어 있는 우리나라 보건의료체계에서는 신의료기술평가에서 통과되더라도 시장진입 여부를 결정하는 단계에서와 다른 기준으로 건강보험 급여 여부를 결정짓기 때문에 사실상 신의료기술평가를 통해 도출된 근거를 건강보험 급여 의사결정에 반영, 연계하는 것이 제도적으로 미흡할 수밖에 없는 구조를 갖고 있다.

그럼에도 불구하고, 지속적인 신의료기술평가제도 규제완화 정책으로 시장진입이 가능한 의료기술이 확대되고 있는 실정에서 신의료기술평가 진입단계에서의 합리적인 의사결정에 대한 중요성이 강조되고 있고, 신의료기술평가가 보다 건강보험제도 운영에 적극 활용될 수 있도록 연계 강화에 대한 필요성이 증대되고 있다.

신의료기술평가에서 통과된 기술의 건강보험제도권 내 의사결정시 임상적 근거자료로서 지원하면서 의료기술재평가를 연속성 있게 수행하기 위해서는 도입과정시 신의료기술평가 결과에 대해 연계가능한 권고안 개선이 필요하다.

▢ 연구 목적

본 연구는 신의료기술평가제도와 관련하여 여러 제도개선을 통해 변화된 신의료기술평가 방법상의 변화와 의료기술 재평가 정책방향을 고려하여 향후 도입되는 신의료기술에 대한 건강보험 의사결정 지원 및 재평가제도와의 연계를 강화시키기 위해 신의료기술평가 권고체계 개편(안)을 제시하고자 하였다.

▢ 연구 방법

국내외 자료원을 통해, 의료기술평가를 수행하는 주요 기관의 근거수준 평가방법과 권고등급 체계 및 결과활용 등에 대해 고찰하였다.

신의료기술평가제도가 도입된 2007년도 이후 2017년까지 신청된 기술에 대해, 평가가 수행된 사례를 토대로 근거자료의 형태 및 근거수준 특징을 분석하여, 신의료기술평가 의사결정의 근거가 된 현황을 파악하고, 근거수준이 의사결정에 어떠한 영향을 미쳤는지를 확인하였다. 또한, 신의료기술평가 결과가 건강보험권 내에서의 의사결정시에 도움을 주는 근거자료로써 활용된다는 점을 고려하여, 신의료기술평가 결과에 따른 급여현황을 토대로 급여의사결정에 미치는 영향을 분석하여 현 신의료기술평가 권고체계 개선의 필요성에 대한 시사점을 도출하였다.

신의료기술로 인정된 기술 중 의료행위전문평가위원회 회의결과 공개자료를 통해 급여여부결정시 논의되었던 사유를 확인할 수 있는 기술에 대해서는 건강보험심사평가원에서 급여여부에 대한 의사결정시 판단요소를 확인하기 위해 급여종류별로 각 논의되었던 사유에 대해 분석을 수행하였다.

전문가 자문과 국내 신의료기술평가 결과의 현황 분석된 내용을 통해 현재의 근거수준 및 권고등급 체계의 제한점을 확인하였으며, 신의료기술평가 권고내용이 활용되는 측면을 고려하여 급여의사결정에 영향을 미칠 수 있도록 지원할 수 있는 관점에서 도출한 권고등급 평가 체계 방안과 의료기술평가관점에서 보다 근거에 입각한 임상적 권고의 강도를 제시하는 방안(GRADE 방법론 활용방안)으로 권고체계 개선(안)을 마련한 후, 전문가 자문회의를 통해 문제점 및 보완사항에 대해 검토하였다.

연구에서 도출된 권고체계(안)이 현재의 신의료기술평가 체제 안에서 적용 가능할 지, 수정보완 및 개선사항을 도출하여 수용성을 제고하고자, 과거 신의료기술로 인정된 진단검사 및 처치 및 시술 영역별 각 사례를 추출하여 시범 적용을 수행하였다.

▢ 연구 결과

권고체계는 현재의 신의료기술평가제도의 결정구조를 벗어나지 않는 범위내에서, 신의료기술(권고함), 혁신의료기술(조건부권고), 연구단계기술(권고하지 않음)의 3

등급 체계로 분류하였고, 신의료기술에 대한 권고등급은 '가 - 권고강도 높음'과 '나 - 권고강도 낮음'으로 제시하였다.

신의료기술의 권고등급을 결정하기 위한 평가기준은 신의료기술에 대한 급여 의사결정시에 중요하게 고려되었던 판단근거요소를 위주로 한 평가기준(1안)과 GRADE의 근거기반 의사결정 틀(Evidence to Decision framework)을 활용한 평가기준(2안)으로 접근하였다. 평가기준 1안은 신의료기술로 인정 받은 기술은 건강보험 등재시 활용된다는 점을 고려하여 보험등재수준 결정과의 연계 강화 차원의 정부 및 보험자적인 관점에서 접근하였다. 건강보험 요양급여결정과정에서 사용되고 있는 가치와 과거 신의료기술로 인정된 기술에 대해 의료행위전문평가위원회에서 의사결정하였던 사례 분석을 토대로, 문헌적 근거 이외 신의료기술평가 단계에서 보강해야 할 평가요인들을 임상적 유용성, 기술의 확립성, 질병의 중증도/희귀성/질병부담, 대체가능기술유무, 재평가 필요성으로 보았다.

평가기준 2안은 신의료기술이 임상전문가들에 의해 적극 활용된다는 점을 고려하여 의과학적 근거에 대한 가치평가를 토대로 임상적 권고의 강도를 제시할 수 있도록 개선하고자 하였다. GRADE EtD 틀을 활용하되, 신의료술평가가 시장진입여부를 결정하는 단계에서의 평가라는 점을 감안하여, 신평가기술(중재) 관련 질병부담, 효과(기대하는 효과, 기대하지 않는 효과), 효과추정에 대한 신뢰도(근거수준), 가치, 효과간의 비교를 평가요소로 한정하여 살펴보았다.

▢ 결론 및 정책적 제언

본 연구에서는 SIGN의 근거수준 체계를 유지하면서, 신의료기술평가 결과의 급여결정과 의료기술평가 전주기적 관점에서 재평가 대상 선별을 위한 합리적인 의사결정하는데 도움을 주기 위해 신의료기술 권고강도에 대한 등급화 하는 방안을 제시하고자 하였다. 그동안 신의료기술로 승인되어 시장진입된 이후 급여등재시 의사결정되었던 사례 분석과 의료기술평가기관에서 사용되는 체계 등을 참고하여 권고등급을 결정하기 위한 2개의 평가기준(안)을 제시하고, 이에 따른 권고등급 체계를 제안하였다.

신의료기술평가 보고서의 평가결과 등급을 근거수준 이외의 다른 요인들을 고려한 권고등급안을 제시하였으나, 이를 신의료기술평가 과정에서 적용하기 위해서는

아직 더 검토해야 할 요인들이 많이 있다. 신의료기술평가에 보다 적합한 평가기준에 대해 전문가 및 이해관계자의 의견수렴을 통해 결정해야 하고, 결정된 평가기준(항목)에 대한 많은 시범적용을 통해 타당도 검증이 이루어져야 사용될 수 있을 것이다. 또한, 본 연구진에서는 신의료기술로 인정된 경우에만 권고강도를 제시하는 것으로 마련하였으나, 실제 안전성 유효성 평가 소위원회를 운영하면서 권고의 강도에 대한 평가가 이루질 것을 감안하면 신의료기술로 인정되는 기술 뿐 아니라 연구단계기술에 있어서도 권고의 강도가 제시될 필요성에 대해서도 추후 검토가 이루어져야 할 것이다.

동 평가내용은 기존의 신의료기술 문헌적 근거 평가 이외 다양한 다른 가치 항목들을 포함하여 평가가 필요하므로, 신의료기술평가 법정기한이 단축되어진 상황에서 평가를 위한 시간적, 인력적 자원이 부족하여 이를 적용하기 위해서는 제도적 보완이 필요하다.

그동안 신의료기술평가는 지속적인 규제완화 정책으로 체계적문헌고찰을 수행하지 않는 신속문헌고찰을 통해 시장진입의 기회를 확대하였고, 건강보험 보장성 강화로 보험권 내에서의 합리적인 의료기술의 적정사용을 유도하기 위해서는 신의료기술평가 단계에서의 의료기술에 대한 합리적인 의사결정이 요구되는 실정에서, 이 연구는 기존의 평가체계를 바탕으로 개선안 마련을 위한 기반 틀을 제시하였다는 점에 의의가 있다. 본 연구를 바탕으로, 후속 연구를 통해 보다 안정적이고, 견고한 방법론으로 수행하여 신의료기술평가 권고 체계를 안착시키는 것이 필요할 것이다.

주요어

신의료기술평가, 권고등급, 근거수준, 근거에 기반한 의사결정 틀

I 서론

1. 연구배경 및 필요성

1.1. 신의료기술평가제도

우리나라에서 신의료기술평가의 대상은 관련 법령에서 안전성·유효성이 평가되지 아니한 의료기술로 보건복지부 장관이 평가가 필요하다고 인정한 의료기술로 제시되어 있다[1]. 그러나 실제적으로는 국민건강보험 요양급여의 기준에 관한 규칙에 따라 요양급여대상 또는 비급여대상으로 결정되지 아니한 새로운 행위 및 새로운 치료재료를 사용하는 행위로 급여 결정이 되지 않은 모든 의료기술을 지칭하고 있다[2]. 신의료기술의 평가는 의료법에 신의료기술평가위원회의 심의를 거쳐 안전성·유효성 등에 관한 평가를 하도록 되어 있다[3]

1.2. 신의료기술평가방법

가. 신의료기술평가 방법론

신의료기술평가는 체계적이고 객관적인 평가를 위하여 해당 의료기술에 대한 임상적 문헌근거를 체계적으로 검색, 분석, 합성 및 요약하여 과학적인 근거를 찾는 연구방법론인 체계적 문헌고찰(systematic review; SR)을 주로 활용하여 수행한다.

체계적문헌고찰은 특정 연구 질문에 대해 최선의 가용 가능한 연구 결과를 종합하는 연구 방법으로써, 체계적이고 포괄적인 문헌검색과 사전에 정해진 포함/배제

1) 신의료기술평가에 관한 규칙 제2조제1항
2) 박은철. 국민 안전과 의료기술 조기활용을 위한 신의료기술평가 제도개선 방안 연구. 2016.
3) 의료법 제53조제1항

기준에 따른 문헌선택, 선정된 문헌에 대한 비뚤림 위험 평가 등의 엄격하고 객관적인 연구 과정을 거친다. 체계적 문헌고찰은 다른 연구와 마찬가지로 해당 분야의 임상전문가, 방법론 전문가, 통계학자 등의 다양한 전문가가 협력해야 하는 다학제적인(multi-disciplinary) 연구방법론이다.[4] 이외에도, 해당 의료기술의 성격 또는 평가대상에 따라 환자 등 설문조사, 원자료 분석, 교과서 및 가이드라인 검토 등의 평가방법을 보완적으로 활용하여 평가가 이루어진다.

나. 신의료기술평가의 절차

신의료기술평가의 평가절차는 처치 및 시술 및 기타검사를 포함한 일반 평가와 체외진단검사(유전자검사 포함) 평가로 구분되어 진행된다.

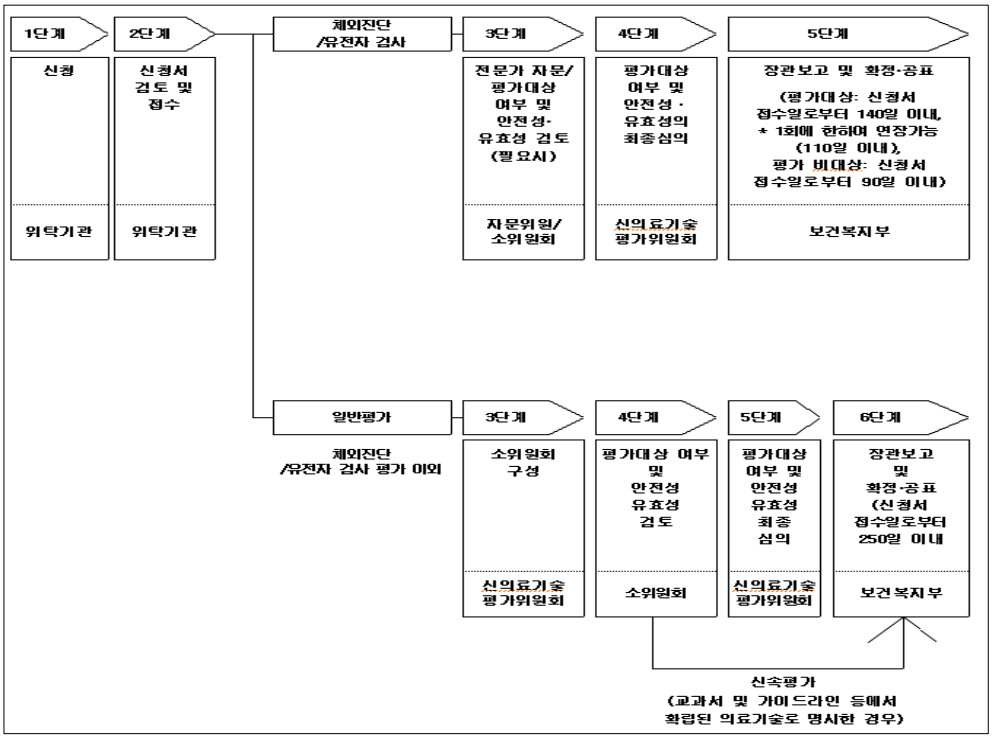

그림 1. 신의료기술평가 절차
(출처 : 신의료기술평가의 절차와 방법에 관한 규정, 고시 2019.3.15.)

4) NECA 체계적문헌고찰 매뉴얼

1.3. 신의료기술평가 결과

신의료기술의 평가결과는 크게 '기존기술(건강보험 급여 또는 비급여 대상과 같거나 유사한 경우)', '안전성·유효성이 있는 의료기술(신의료기술)'과 '연구단계 의료기술(안전성 또는 유효성이 확인되지 아니한 의료기술)'로 구분될 수 있다[5]. 신의료기술평가의 결과를 통해, 안전하고 유효한 기술로 인정된 경우 요양급여결정행위 신청을 통해 의료현장으로의 시장진입이 가능하다.

체계적문헌고찰 방법을 수행하여 근거자료를 마련한 경우에는 영국 Scottish Intercollegiate Guidelines Network (SIGN)[6]의 질평가 판정기준(표 2)을 적용하여 해당 의료기술에 대한 문헌적 근거수준(근거강도)(표 3, 표 4)을 평가결과와 함께 제시하고 있다. 참고적으로, 표4는 SIGN에서 권고의 등급(Grade of Recommendation)으로 제시되고 있으며, 과거 신의료기술평가 보고서에서도 SIGN의 사용된 용어를 원용하여 권고등급으로 표기하였다. 하지만 이는 연구유형에 따른 질평가 결과를 종합한 근거의 강도와 관련된 내용으로써, 임상적 중요성을 반영하는 것은 아니었으나 이 의미를 임상적 권고로 잘못 혼동된 경우가 빈번히 발생하여 2016년 12월부터는 '권고의 등급'을 '근거의 수준'으로 수정하여 제시하고 있으며 현재 신의료기술평가에서 별도의 권고등급은 제시하지 않고 있다.

다만, 연구가 더 필요한 단계의 기술로 심의될 경우 해당 기술의 잠재적 이익 등을 고려하여 2011년부터는 등급분류[7]를 시작하였으며(표 5 참고), 2014년부터는 제한적 의료기술 제도를 도입하여 연구단계기술 중 안전성이 확보된 의료기술로서 대체기술이 없는 질환이거나 희귀질환의 치료·검사에 대해 의료기관의 신청을 받아 신의료기술평가위원회의 심의를 거쳐 선정된 의료기관에서 비급여 진료를 허용하고 그 결과를 의과학적 근거로서 활용할 수 있게 하였다[8]. 제한적 의료기술은 2016년 11월에 그 대상범위를 확대하여 현재는 연구단계기술로 심의된 기술 중 안전성이 확보되었고, 임상에 도입할 필요가 있는 의료기술의 비급여 진료 조건부 허용을 통해 근거창출의 기회를 제공할 수 있도록 하고 있으며[9], 신청대상 후보기술은 질환의 중증도 및 잠재적 영향력 등을 고려하여 등급분류를 하고 있다(표 6참고).

---

5) 신의료기술평가에 관한 규칙 제3조제8항
6) SIGN 홈페이지. available from URL: http://www.sign.ac.uk/
7) 신의료기술평가의 절차와 방법 등에 관한 규정 별표 2(보건복지부고시 제2011-80호)
8) 신의료기술평가에 관한 규칙 제3조제8항제2호(보건복지부령 제236호, 2014.4.24., 일부개정)
9) 신의료기술평가에 관한 규칙 제3조제8항제2호(보건복지부령 제444호, 2016.11.14., 일부개정)

표 1. 신의료기술평가결과

심의결과	설명	
신의료기술	해당 평가대상 의료기술의 안전성 및 유효성이 인정되어 임상에서 사용가능한 의료기술	
연구단계 의료기술	I등급	임상 도입시 잠재적 이익이 크지 않은 경우
	II-a 등급	대체 기술은 존재하나, 임상 도입 시 잠재적 이익이 큰 의료기술로 임상지원이 필요하다고 판단되는 경우
	II-b 등급	대체 기술이 없는 의료기술, 또는 희귀질환 치료(검사) 방법으로 남용의 소지가 없는 의료기술로 시급한 의료현장도입이 필요하다고 판단되는 경우

표 2. 질 평가 판정기준

++	거의 모든 또는 모든 기준이 충족됨. 연구나 검토의 미 충족된 부분으로 인해 연구의 결론이 바뀌지 않을 것으로 확신될 경우
+	몇 가지 기준이 충족됨. 부적절하거나 미 충족된 부분으로 인해 연구의 결론이 바뀌지 않을 것으로 생각되는 경우
−	거의 모든 또는 모든 기준이 충족되지 않음. 이 연구의 결론은 바뀔 것으로 생각되는 경우

표 3. 질 평가 결과

1++	• 무작위 임상시험으로 수행된 높은 질의 메타분석과 체계적 문헌고찰 • 비뚤림의 위험이 매우 낮은 무작위 임상시험
1+	• 무작위 임상시험 또는 비무작위 임상시험으로 잘 수행된 메타분석과 체계적 문헌고찰 • 비뚤림의 위험이 낮은 무작위 임상시험 또는 비무작위 임상시험
1−	• 무작위 임상시험 또는 비무작위 임상시험으로 수행된 메타분석과 체계적 문헌고찰 • 비뚤림의 위험이 높은 무작위 임상시험 또는 비무작위 임상시험
2++	• 환자−대조군 또는 코호트 연구 및 진단법 평가연구로 수행된 높은 질의 체계적 문헌 고찰 • 교란(confounding)이나 비뚤림 및 우연성의 위험이 매우 낮거나 인과 관계에 대한 높은 확률을 가진 높은 질의 환자−대조군 또는 코호트 연구 및 진단법 평가연구
2+	• 교란이나 비뚤림 및 우연성의 위험이 낮거나 인과 관계에 대한 보통의 확률을 가진 높은 질의 환자−대조군 또는 코호트 연구 및 진단법 평가연구
2−	• 교란이나 비뚤림 및 우연성의 위험이 매우 높거나 인과 관계가 없는 상당한 위험을 환자−대조군 또는 코호트 연구 및 진단법 평가연구

3	• 비분석적 연구 (예: 전후 연구, 증례 연구, 증례 보고)
4	• 전문가 의견

표 4. 근거의 수준

근거수준	정의
A	• 메타분석 및 체계적 문헌고찰 또는 1++의 무작위 임상시험 연구가 최소 하나 이상이고, 표적 모집단에 직접 적용 가능한 경우
	• 주로 1+의 연구로 구성된 근거이고, 직접 표적 모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
B	• 2++의 연구로 구성된 근거이고, 직접 표적 모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
	• 1++나 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 근거인 경우

표 5. 조기기술 및 연구단계 의료기술 등급 부여 기준

분류		기준
Ⅰ등급		임상도입시 잠재적 이익이 크지 않은 경우
Ⅱ등급	Ⅱ-a등급	대체기술은 존재하나, 임상도입시 잠재적 이익이 큰 의료기술로 임상지원이 필요하다고 판단되는 경우
	Ⅱ-b등급	대체기술이 없는 의료기술, 또는 희귀질환 치료(검사)방법으로 남용의 소지가 없는 의료기술로 임상지원이나 시급한 임상도입이 필요하다고 판단되는 경우

표 6. 제한적 의료기술 분류 기준

분류	기준	비고
A	1. 질환의 중증도가 높은 환자를 대상으로 한 의료기술 2. 기술의 난이도가 높거나 시술(검사)시 위험도가 높은 의료기술 3. 의료결과에 있어 상당한 영향력이 예상되는 의료기술	
B	1. 질환의 중증도가 높지 않은 환자를 대상으로 한 의료기술 2. 기술의 난이도가 높지 않고 시술(검사)시 위험도가 높지 않은 의료기술 3. 의료결과에 있어 잠재적 영향력이 예상되는 의료기술	

분류	기준	비고
C	1. 안전성에 문제가 있거나 확인되지 않은 의료기술 2. 임상적 유용성이 없거나 현저히 떨어지는 의료기술	제한적 의료기술대상 제외

출처 : 신의료기술평가사업본부 홈페이지_「제한적 의료기술」 (대상영역 확대) 신청공고. 2017.10.30.기준

1.4. 신의료기술평가 관련 제도개선 추진현황

가. 신의료기술평가 절차 신속화 및 간소화

신의료기술평가 제도는 의료기술의 조기 시장진입을 위하여 여러가지 제도개선이 진행되어왔다. 2016년부터는 신의료기술의 평가기간을 단축시키기 위해 의료기기의 식품의약품안전처 허가와 신의료기술평가를 통합운영 및 검사분야에 있어 신의료기술평가 제외대상을 확대하였으며, 2017년도에는 신의료기술평가방법에 있어 중재시술과 기타검사 영역에서 체계적 문헌고찰 방법 이외에 교과서 및 가이드라인을 통해 확립된 기술로 확인되면 신속평가를 수행하여 신의료기술로 심의할 수 있도록 심의기준을 개정하였다.

표 7. 중재시술의 평가대상여부 심의 기준

의료기술 유형	심의 기준
1. 의학교과서 및 가이드라인에서 안전성 · 유효성이 확립된 기술이지만 별도 행위로 분류되어 있지 않은 경우	신의료기술(신속평가)
2. 새로운 기술	평가대상
3. 사용 대상이 변경된 경우	평가대상
4. 사용 목적이 변경된 경우	평가대상
5. 사용방법이 변경된 경우	
5-1. 수술 경로가 변경된 경우	평가대상
5-2. 외과적 수술에서 중재적 기법으로 변경된 경우	평가대상
5-3. 수술 절차가 변경된 경우	평가대상
5-4. 에너지원이 변경된 경우	평가대상(신속평가)
5-5. 레이저 종류가 변경된 경우	기존기술
5-6. 기존 시술에 시술방법이 추가된 경우	기존기술

의료기술 유형	심의 기준
5-7. 수기요법이 자동으로 변경된 경우	기존기술
5-8. 치료재료가 변경된 경우	기존기술/ 평가대상(신속평가)

표 8. 기타 검사의 평가대상여부 심의기준

의료기술 유형	심의 기준
1. 의학교과서 및 가이드라인에서 안전성 · 유효성이 확립된 기술이지만 별도 행위로 분류되어 있지 않은 경우	신의료기술(신속평가)
2. 새로운 검사	평가대상
3. 사용 대상이 변경된 경우	평가대상
4. 사용 목적이 변경된 경우	평가대상
5. 사용 방법이 변경된 경우	평가대상

나. 혁신의료기술평가 별도트랙 신설

신의료기술평가는 그동안 임상문헌 중심으로 안전성과 유효성을 평가함에 따라 개발이력이 짧아 근거문헌이 부족한 기술에 대해서는 시장진입이 지체되어 의료산업의 발전을 저해하는 결과를 초래해 문헌중심의 평가를 보완할 필요성이 제기되었다. 이에, 잠재가치가 높은 혁신의료기술의 가치를 적절하게 평가하여 조기시장진입을 촉진시키기 위한 목적으로 평가방법 및 절차 등을 새롭게 규정하기 위해, 2017년부터 ‘신개발 의료기기 시장진입 활성화를 위한 제도개선 방안 연구’, 2018년 ‘혁신의료기술 도입을 위한 별도 평가 체계 구축에 관한 연구’를 진행하였으며, 2018년에는 정부에서 의료기기 규제 혁신 대책을(‘18.7.1.9) 발표를 통하여, 혁신의료기술 별도 평가 트랙을 마련하도록 하였다.

혁신의료기술 별도 평가 트랙은 인공지능, 3D 프린팅, 로봇 등의 첨단 기술이 접목된 혁신의료기술에 대해 문헌기반의 안전성 유효성 평가와 더불어 기술의 혁신성, 대체기술 유무, 질병의 심각성, 남용가능성, 환자 삶의 질 향상 정도 등의 포괄적 가치 평가척도(표 9)를 포함하여 함께 평가하는 것이며, 제도가 안정적으로 안착될 수 있도록 시범사업을 2018년 9월부터 ‘19년 1월까지 진행하였다.

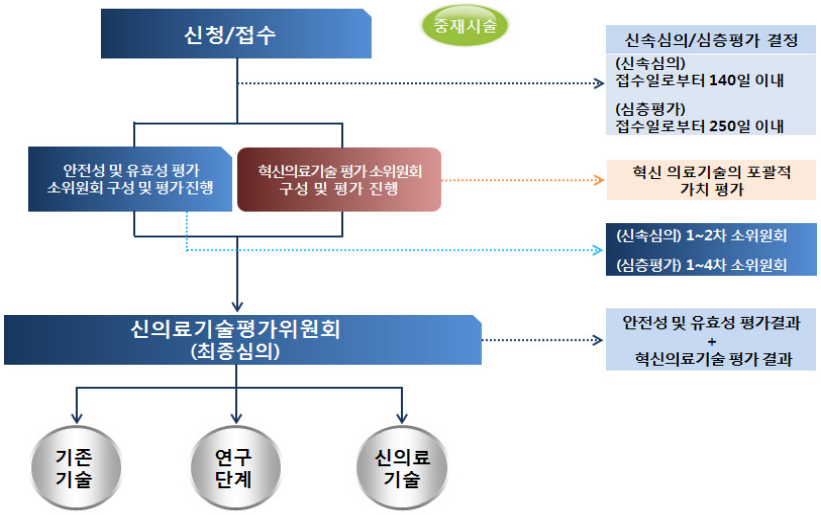

그림 2. 혁신의료기술 별도 평가 절차

표 9. 혁신의료기술 잠재성 평가 척도

평가 항목		평가 기준
질병	질병의 중요성	해당질환의 진행과정에서 사망, 합병증, 장애 및 후유증 등에 미치는 영향
	질병의 희귀성	해당 질환이 희귀질환에 해당하는지
환자 중심성	환자의 신체적 부담	기존기술과 비교하였을 때, 동 기술이 환자에게 주는 신체적 부담이나 위해 정도
	환자의 경제적 부담	대체 가능한 기존기술과 비교하였을 때, 환자 본인 부담은 어느 정도인지
	삶의 질	기존기술과 비교하였을 때, 동 기술에 의한 환자의 삶의 질 향상의 정도
사회	남용 가능성	해당 기술이 의료현장에서 남용 가능성이 어느 정도인지
기술	대체기술 유무	해당 기술과 동등 수준 이상의 대체기술이 있는지
	기술의 혁신성	의료현장에 도입 시 임상적 유용성 및 사회적 파급력이 어느 정도인지

혁신의료기술 별도 평가 트랙을 통한 신의료기술의 평가결과는 일반 절차와 마찬가지로 크게 안전성·유효성이 있는 기술(신의료기술)과 연구가 더 필요한 단계의 기술로 구분될 수 있다. 다만, 신의료기술로 인정되는 경우는 안전성에 대한 우려는 적으나 문헌적 근거로서 임상적 유효성이 불확실한 의료기술에 대해 잠재성의 가치를 인정받은 것이기 때문에 오남용의 방지를 위해 특정 실시기관에서 일정기간 동안 근거축적을 조건으로 의료현장에서 사용가능한 '조건부 신의료기술'이라는 점에서 일반적으로 승인되는 신의료기술과는 차이가 있다. '조건부 신의료기술'은 연구과정에 대한 보고 및 모니터링을 실시하고, 일정기간(3~5년)이 경과 후 창출된 근거를 포함하여 근거 기반의 재평가를 실시하게 된다.

다. 체외진단 先(선)-진입 後(후)-평가

체외진단검사는 인체 내 채취된 혈액 등의 검체를 통해 질병의 진단 및 예후예측, 관찰 등을 목적으로 체외에서 수행되는 검사를 의미한다.

2018년 정부는 의료기기 규제 혁신 대책을('18.7.1.9) 발표하여, 인체에 직접 사용되지 않아 안전성 우려가 적은 체외진단검사 분야에 대해 식약처 의료기기 허가를 득한 이후 신의료기술평가 절차를 유예하여 바로 의료시장에 진입을 허용하게 하고, 일정 기간이 지난 후에 신의료기술평가를 수행하겠다고 규제완화 계획을 발표하였다. 기존에 이미 운영되고 있는 평가유예 신의료기술평가제도(임상시험을 거쳐 식약처 허가를 받은 신의료기기를 사용한 의료행위에 대해 신의료기술평가를 1년간 유예하여 조기에 임상현장에서 활용)를 체외진단검사분야로 확대시킨 것으로 볼 수 있다.

선-진입 대상이 되는 체외진단검사 의료행위 기술은 식약처 허가에서 사용목적이 검사대상, 분석물질, 검사질환 등이 명시되어 임상적 목적이 특정된 체외진단용 의료기기를 수반한 의료기술이 해당된다. 2019년 상반기에는 우선적으로 감염병 체외진단검사분야에 한해 시범운영 한 후에, 하반기에 단계적으로 대상의 범위를 확대할 계획에 있다.

체외진단검사가 인체에 직접적 위해가 낮아 비교적 검사수행으로 인한 안전성의 문제가 적다고 하지만, 일부 유효성이 떨어지는 체외진단검사의 시장진입으로 인한 오남용 발생의 문제가 존재하는 바, 시장진입 후 사용기간 동안 근거창출 수행에 대한 과정관리 및 실시기관을 제한하는 등 관리기전 마련에 대해 검토하고 있다. 또한, 평가 유예기간동안에 신의료기술평가 신청을 의무 부여하여, 신의료기술 후평가를 통하여 심평원 급여의사결정에 활용할 수 있도록 지원할 계획에 있다.

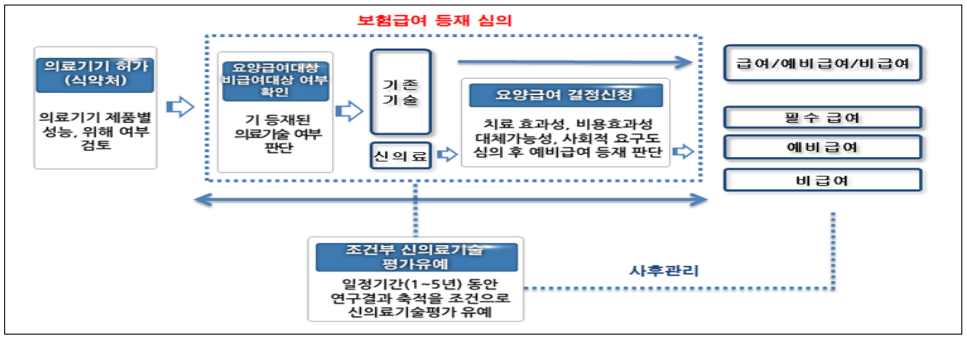

그림 3. 체외진단검사 등재 절차 개선(안)

1.5. 의료기술평가 재평가

우리나라는 신의료기술평가제도를 통해 안전하고 유효한 새로운 의료행위의 도입을 관리하는 체계가 마련되어 있지만, 이미 건강보험권에 진입된 이후 사용되고 있는 기존의료기술에 대해서는 그동안 관리가 부재한 상황이었다. 신의료기술평가의 지속적인 제도개선으로 새로운 의료기술의 시장진입이 확대되는 상황과, 2017년 정부에서 발표된 건강보험 보장성 강화 정책에 따라 의료기술의 전주기적 관리 차원에서 기존에 사용되고 있는 의료기술에 대한 재평가 필요성이 강조되고 있다. 이와 관련하여, 2018년에 보건복지부가 의료기술평가를 위한 기본 계획을 수립하여 발표하였으며, 한국보건의료연구원에서 실제 재평가 운영이 수행 가능하도록 세부운영 절차 등의 체계구축 연구를 진행하였다. 또한, 2018년에 건강보험심사평가원으로부터 의뢰받은 비급여 항목 중 일부 의료기술에 대해 의료기술재평가를 수행하였으며, 2019년에도 건강보험심사평가원으로부터 의뢰받은 비급여 항목 및 신료기술평가 업데이트 재평가를 지속적으로 수행 중에 있다.

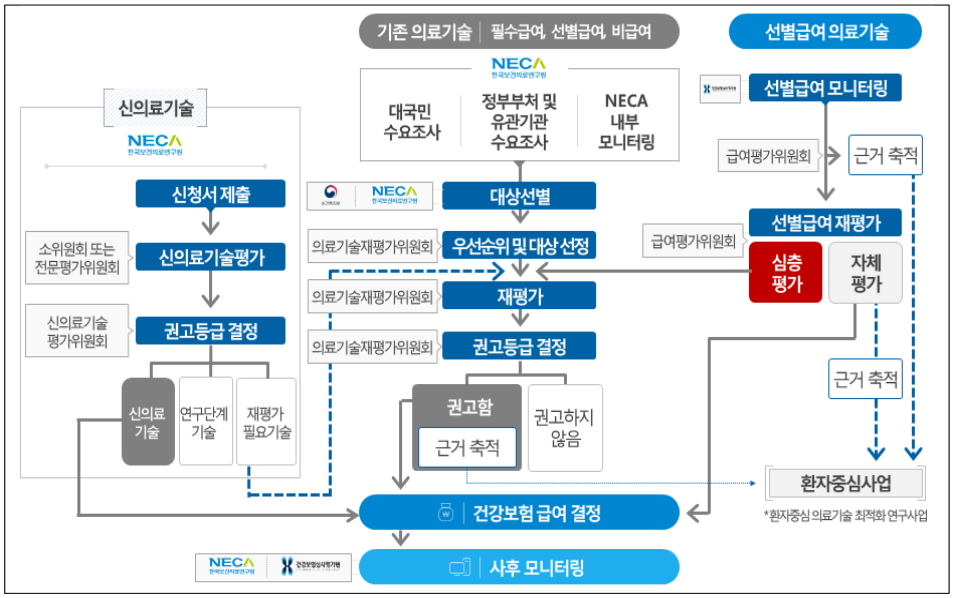

그림 4. NECA의 의료기술재평가 기본체계도(안)
(출처 : 박동아 등, 의료기술 재평가체계 구축 및 운영방안 연구. 2018)

현재의 보건의료체계에서는 신의료기술평가 결과를 통해 안전하고 유효하다고 통과된 의료기술은 모두 향후 의료기술재평가 대상이 되므로, 신의료기술평가와 재평가와의 연계가 원활하게 작동될 수 있도록 신의료기술평가 권고가 제시될 필요가 있다.

1.6. 연구의 필요성

우리나라는 2007년 신의료기술평가 제도가 도입되면서 새로운 의료기술의 시장 도입 기전이 확립되어 의료서비스에서 사용되는 의료기술의 임상 도입, 급여화 등의 관리를 국가가 책임지도록 규정함으로써 모든 의료기술을 건강보험 급여 체계와 연계하도록 제도화되었다[10].

신의료기술평가는 기본적으로 체계적 문헌고찰을 통해 문헌적 근거에 기반한 의학적 근거 수준을 평가하여 의료기술의 도입에 관한 의사결정과정을 수행하고 있

10) 박종연. 보건복지포럼. 의료기술의 평가와 급여 결정방향. 2017

고, 신의료기술평가 결과는 안전하고 유효한 '신의료기술'과 '연구단계기술'로 크게 구분되어 평가되고, 체계적문헌고찰을 수행한 경우에는 SIGN의 질평가 결과에 따른 문헌적 근거수준(A~ D)이 함께 제시되고 있다.

신의료기술평가에서 체계적 문헌고찰시 사용되고 있는 SIGN의 질 평가 도구는 연구의 설계에 따른 잠재적 비뚤림에 대한 가능성을 평가한다. SIGN 평가로부터 도출된 등급은 그동안 신의료기술평가시 임상적 권고의 강도로 오해해 등급이 A로 높을수록 보험 등재시 급여 결정에 긍정적인 영향을 미칠 것으로 잘못 이해되기도 하였다. 연구설계가 반드시 임상적 중요성을 반영하지 않음에도 불구하고 SIGN의 평가등급은 연구설계에 따른 연구의 질을 기준으로 일관성, 적용가능성을 고려하여 결정된 등급으로써 근거에 대한 효과의 크기를 평가하지 않는 등 연구설계에 따른 근거수준이 직접적으로 반영된 등급이라는 제한점이 있었다.

신의료기술평가에서 신의료기술로 인정되어야 의료행위의 건강보험 급여 신청이 가능하고, 신의료기술의 평가결과는 건강보험과정안에서의 급여기준마련의 근거자료로 활용되고 있다. 건강보험심사평가원에서의 의료행위의 급여여부는 의학적 타당성, 의료적 중대성, 치료효과성, 비용효과성, 환자의 비용부담정도 및 사회적 편익 등을 고려하여 결정하도록 규정되어 있다. 한국보건사회연구원회의 연구보고서[11]에 따르면, 2007년~2011년 동안 '신의료기술'로 인정받아 건강보험 급여 신청을 하여 심의된 의료행위 488건 중 116건(23.6%)이 급여, 372건(76.2%)이 비급여로 결정되었다고 보고된 바 있다. 이렇게 비급여 비율이 높은 것은 임상적 유용성 및 비용효과성의 기준을 충족하지 못하는 것이 가장 큰 이유로 제시되고 있다.

이처럼, 현재 신의료기술의 도입에 있어 신의료기술평가단계와 건강보험 결정 단계로 이분화되어 있는 우리나라 보건의료체계에서는 신의료기술평가에서 통과되더라도 시장진입 여부를 결정하는 단계에서와 다른 기준으로 건강보험 급여 여부를 결정짓기 때문에 사실상 신의료기술평가를 통해 도출된 근거를 건강보험 급여 의사결정에 반영, 연계하는 것이 제도적으로 미흡할 수밖에 없는 구조를 갖고 있다.

그럼에도, 지속적인 신의료기술평가제도 규제완화 정책으로 시장진입이 가능한 의료기술이 확대되고 있는 실정에서 신의료기술평가의 진입단계에서의 합리적인 의사결정에 대한 중요성이 강조되고 있으며, 신의료기술평가가 보다 건강보험제도 운영에 적극 활용될 수 있도록 연계 강화에 대한 필요성이 증대되고 있다.

또한, 신의료기술평가는 임상에서 사용되기 전의 제한된 근거로 도입이 결정된

11) 박실비아 등. 신의료기술의 도입과 확산에 관한 정책분석. 한국보건사회연구원. 2014

것이기 때문에 의료기술의 시장도입 이후에 의료현장에서 사용되면서 축적된 안전성·유효성, 비용효과성 등에 대한 근거는 사회적 환경 및 평가시점에 따라 달라질 수 있다. 한정된 보건의료자원하에 보다 효율적인 의료기술의 사용을 권장하기 위해서는 기존 의료기술에 대한 재평가는 중요하다. 정책적으로도 의료기술의 재평가의 중요성을 인지하고, 의료기술평가를 위한 기본 계획을 수립한 상황에서 앞으로 평가되는 신의료기술 인정된 기술은 재평가의 대상이므로 신의료기술평가와 재평가와의 연계가 원활하게 작동될 수 있도록 신의료기술평가 권고가 제시될 필요가 있다.

따라서, 신의료기술평가에서 통과된 기술의 건강보험제도권 내 의사결정시 임상적 근거자료로서 지원하면서 의료기술재평가를 연속성 있게 수행하기 위해서는 도입과정시 신의료기술평가 결과에 대해 연계가능한 권고안 개선이 필요하다.

2. 연구목적

본 연구는 신의료기술평가제도와 관련하여 여러 제도개선을 통해 변화된 신의료기술평가 방법상의 변화와 의료기술 재평가 정책방향을 고려하여 향후 도입되는 신의료기술에 대한 건강보험 의사결정 지원 및 재평가제도와의 연계를 강화시키기 위해 신의료기술평가 권고체계 개편(안)을 제시하고자 하였다. 이에 구체적인 목적은 다음과 같다.

첫째, 주요국의 의료기술평가 및 진료지침의 근거수준 및 권고등급 체계 현황을 조사하여 파악한다.

둘째, 신의료기술평가 결과 현황을 분석을 통해 기술별 근거수준 및 특징을 파악하여 현 신의료기술평가 권고체계에 대한 문제점을 도출한다.

셋째, 신의료기술평가 인정 기술에 대한 급여의사결정 사례를 분석하여, 연계강화를 위해 필요한 권고평가시 고려항목을 도출한다.

넷째, 신의료기술평가결과의 권고체계(안)을 마련한다.

다섯째, 개편된 ‘권고체계’을 바탕으로 과거 신의료기술 인정된 기술에 시범적용을 수행한다.

Ⅱ 연구방법

국내외 자료원을 통해, 의료기술평가를 수행하는 주요 기관의 근거수준 평가방법과 권고등급 체계 및 결과활용 등에 대해 고찰하였다. 또한, 신의료기술평가는 급여결정과정에서 의료기술평가를 수행하는 외국의 체계와는 다르게, 국내 식약처 허가를 받은 의료기술 중 보험권에 진입가능한 의료기술을 결정하는 즉, 시장진입을 결정하는데 활용되고 있는 점을 고려하여, 시장진입단계에서 평가되고 있는 국외 체계를 검토하였다.

신의료기술평가제도가 도입된 2007년도 이후 2017년까지 신청된 기술에 대해, 평가가 수행된 사례를 토대로 근거자료의 형태 및 근거수준 특징을 분석하여, 신의료기술평가 의사결정의 근거가 된 현황을 파악하고, 근거수준이 의사결정에 어떠한 영향을 미쳤는지를 확인하였다. 또한, 신의료기술평가 결과가 건강보험권 내에서의 의사결정시에 도움을 주는 근거자료로써 활용된다는 점을 고려하여, 신의료기술평가 결과에 따른 급여현황을 토대로 급여의사결정에 미치는 영향을 분석하여 현 신의료기술평가 권고체계 개선의 필요성에 대한 시사점을 도출하였다.

신의료기술평가 사례별 결과는 여러 가지 적응증 및 시술방법을 포함하여 각각의 평가결과 및 권고등급을 제시한 사항을 고려하여 자료추출을 수행하였다. 자료추출은 먼저 서식안을 마련하였으며, 시범으로 자료추출을 수행한 이후 자문회의를 통해 수정된 자료추출 서식(안)을 이용하여 수행하였다. 결과 분석은 평가결과별, 기술유형별 근거수준 비율을 확인하는 방향으로 진행하였다. 신의료기술평가 결과분석은 신의료기술평가 수행의 단계를 고려하여, 신청단계에서의 전체 현황과 신의료기술 평가대상으로 심의되어 심층평가가 수행된 기술들의 현황, 심층평가 후 신의료기술로 인정된 기술된 기술들의 현황으로 구분하여 수행하였으며 각 단계별로 분석 요인은 내부 연구진 회의 및 자문회의를 통하여 적절한 분석 요인을 선정하였다.

신의료기술로 인정된 기술 중 의료행위전문평가위원회 회의결과 공개자료를 통해 급여여부결정시 논의되었던 사유를 확인할 수 있는 기술에 대해서는 건강보험심사평가원에서 급여여부에 대한 의사결정시 판단요소를 확인하기 위해 급여종류별로 각 논의되었던 사유에 대해 분석을 수행하였다.

신의료기술평가의 권고체계 개선(안) 마련에 대한 의견을 청취 및 수렴하기 위해 신의료기술평가 수행 및 심의·의결과정에 참여한 경험이 있는 의료전문가 및 신의료기술평가위원회 전·현직 위원을 대상으로 하여 전문가 자문조사를 실시하여, 현재 체계의 장애요인, 개선필요성 및 요구사항을 파악하였다.

전문가 자문과 국내 신의료기술평가 결과의 현황 분석된 내용을 통해 현재의 근거수준 및 권고등급 체계의 제한점을 확인하였으며, 신의료기술평가 권고내용이 활용되는 측면을 고려하여 급여의사결정에 영향을 미칠 수 있도록 지원할 수 있는 관점에서 도출한 권고등급 평가 체계 방안과 의료기술평가관점에서 보다 근거에 입각한 임상적 권고의 강도를 제시하는 방안(GRADE 방법론 활용방안)으로 권고체계 개선(안)을 마련한 후, 전문가 자문회의를 통해 문제점 및 보완사항에 대해 검토하였다.

연구에서 도출된 권고체계(안)이 현재의 신의료기술평가 체제 안에서 적용 가능할 지, 수정보완 및 개선사항을 도출하여 수용성을 제고하고자, 과거 신의료기술로 인정된 진단검사 및 처치 및 시술 영역별 각 사례를 추출하여 시범 적용을 수행하였다.

Ⅲ 연구결과

1. 주요국 및 기관의 근거수준·권고등급 체계

1.1. 영국

1.1.1. 보건의료체계

○ 영국 국가보건의료서비스(National Health Service, 이하 NHS)는 일반 재정을 재원으로 삼는 보건의료체계로는 그 규모가 세계에서 가장 크다. NHS는 '모든 국민은 지불능력에 상관없이 의료 욕구에 따른 의료서비스를 받을 수 있어야 한다'는 설립 원칙 아래 오늘에 이르기까지 무상 의료를 실현하고 있으며, 스코틀랜드, 북아일랜드, 잉글랜드, 웨일즈는 모두 독자적으로 NHS를 운영하고 있다.

○ 보건의료정책은 국가수준에서 보건부에 의해 결정되나 지역수준에 대한 관점은 지역별로 10개 설립된 Strategic Health Authorities(이하 SHA)에 의해 수행되며, 영국 의료서비스 중 90%가 1차 의료기관에서 처리되고, 환자가 부담하는 비급여 서비스는 거의 없다(김수경 등 2014).

1.1.2. 의료기술평가제도

가, HTA 개요

○ 영국 전역의 의료기술평가가 단일 적용되지 않으며, 지역별로 다르게 운영된다. 대표적인 기관이 National Institute for Health and Care Excellence(이하 NICE)로 잉글랜드와 웨일즈 지방에서만 적용되는 권고안을 제공한다. NICE에서는 의료기술평가 주제 선정 및 의료기술 판정을 담당하며, NHIR Evaluation, Trials and Studies Coordinating Center(이하 NETSCC)에서 의료기술평가를 담당한다.

○ NETSCC에서 NIHR HTA, SR 프로그램을 통해 평가를 수행하며, 정책의사결정 또는 NHS를 위한 지침작성 지원 목적으로 한다. NICE의 사정과정에서는 해당 의료기술의 비용 및 기타 이익을 고려하여 다음 6개 기준을 고려하고 있다.

1. 질병으로 인한 부담
2. 자원에 미치는 영향
3. 임상적이고 정책적인 중요성
4. 실행함에 있어 부적절한 변화가 존재할 가능성
5. 권고 생산을 위한 기한에 영향을 미칠 가능한 요인
6. 권고안이 공중보건과 삶의 질에 영향, 건강불평등 감소 등에 가져올 가능성

○ 평가결과는 NICE 외에도 National Screening Committee와 National Clinical directors and Chief Medical Officer 등에서도 정책결정을 위한 주요 근거자료로 활용되고 있다

나, HTA 수행 기관 및 역할[12)]

가) NIHR (National Institute for Health Research)

○ 영국 보건부 산하 전문연구기관으로 '2006년 4월 설립되었으며, 직접적인 연구수행 외에도 재정지원 및 연구관리, 연구 활성화를 위한 연구시설 및 인력을 제공, 연구수행 및 참여를 위한 개인연구자 지원 그리고 윤리적 연구수행과 연구결과의 효율적 관리 시스템 구축을 통한 임상연구 증진 측면을 기관운영의 핵심으로 하고 있다. 또한, 코크란 평가그룹이 수행하는 보건의료 중재법과 관련된 체계적 문헌고찰 연구를 지원한다.

나) NICE (National Institute for Health and Care Excellence)

○ HTA를 통해 지역별 의료서비스 제공의 차이를 완화하고자 지난 1999년 준정부

12) 국민안전과 의료기술 조기활용을 위한 신의료기술평가 제도개선 방안연구(2016, 박은철)

기관으로 동 기관이 설립되었으며, 엄격하고 독립적이며 객관적으로 수행되는 검증결과를 근거로 1999년부터 NHS에 제공하고 있으며, 2013년 4월 부터는 사회적 돌봄 분야의 권고를 포함하고 있다. 권고 또는 가이드라인은 NHS 시스템 내의 환자가 이용하는 의약품, 의료행위, 의료기기, 진단 등의 효과성 뿐 아니라 경제성(비용효과)이 높은 의료기술에 대한 접근성 제고를 위한 HTA 결과에 근거한 기술사정을 통해 작성된다.

다) NETCC(NIHR Evaluation, Trials and Studies Coordinating Center)

○ 의료와 사회복지 분야 의사결정자들이 요구하는 과학적이며 객관적인 근거 생성을 위해 NIHR에서는 2008년 4월부터 사우스햄튼 대학(the University of Southhampton)을 중심으로 구성된 NETSCC를 통해 NETS(The NIHR, Evaluation, Trials and Studies) 프로그램을 운영하고 있다.

○ 주관기관인 Southhampton의 Southhampton Health Technology Assessment Center(SHTAC)을 포함하여 7개의 대학연구기관과 2개의 전문연구기간 등 총 9개의 연구기관에서 HTA를 수행하고 있다. 각 연구기관은 특성에 따라 의약품, 의료기기, 의료술기, 검진프로그램, 공중보건 등의 의료기술에 대해 근거자료 생성, 근거합성 또는 체계적문헌고찰과 경제성 평가 등의 방법론을 적용하여 NIHR의 의료기술평가 프로그램 수행과 NICE, 보건부의 정책조사 프로그램 등에서 이루어지는 정책의사결정 또는 NHS를 위한 지침작성을 지원한다.

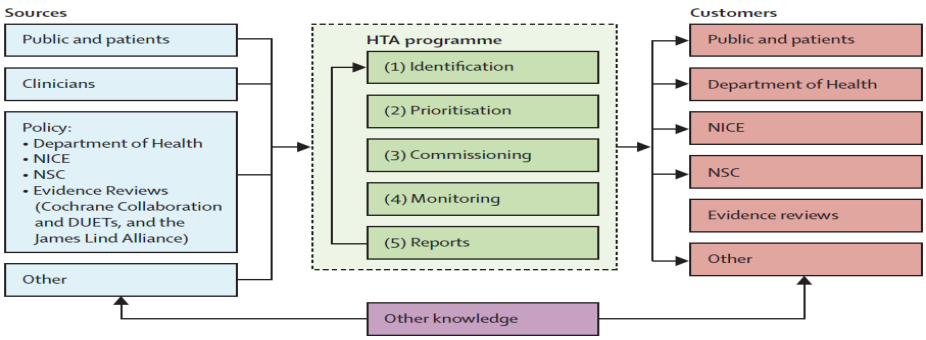

그림 5. The National Institute for Health Research HTA program
(출처 : Raftery J, Powell J. 2013)

다, 근거의 수준[13] 및 권고등급

○ NICE에서는 의료기술평가를 수행 시 별도의 근거수준이나 권고등급을 제시하지는 않고 있다. 다만, SIGN에서 제시하고 있는 근거수준 및 권고등급 체계는 표 10, 표 11과 같으며 연구설계 및 비뚤림 위험, 우연성의 위험, 인과관계 여부에 따라 근거수준을 분류하고 근거수준 및 직접 표적 모집단에 적용 여부와 일관성 여부에 따라 권고등급 체계를 구분해서 제시하고 있다.

표 10. SIGN에서 제시한 근거수준 체계

근거수준	평가 기준
1++	· 무작위임상시험으로 수행된 높은 질의 메타분석과 체계적 문헌고찰 · 비뚤림(bias)의 위험이 매우 낮은 무작위 임상시험
1+	· 무작위임상시험 또는 비무작위임상시험으로 잘 수행된 메타분석과 체계적 문헌고찰 · 비뚤림(bias)의 위험이 낮은 무작위임상시험 또는 비무작위임상시험
1−	· 무작위임상시험 또는 비무작위임상시험으로 수행된 메타분석과 체계적 문헌고찰 · 비뚤림(bias)의 위험이 높은 무작위임상시험 또는 비무작위임상시험
2++	· 환자-대조군 또는 코호트 연구 및 진단법평가연구로 수행된 높은 질의 체계적 문헌고찰 · 혼동(confounding)이나 비뚤림 및 우연성의 위험이 매우 낮거나 인과관계에 대한 높은 확률을 가진 높은 질의 환자-대조군 또는 코호트연구 및 비교연구
2+	· 혼동이나 비뚤림 및 우연성의 위험이 낮거나 인과관계에 대한 보통의 확률을 가진 높은 질의 화낮-대조군 또는 코호트 연구 및 진단법평가연구
2−	· 혼동이나 비뚤림 및 우연성의 위험이 매우 높거나 인과관계가 없는 상당한 위험을 가진 환자-대조군 또는 코호트 연구 및 진단법평가연구
3	· 비분석적 연구(예: 전후연구, 증례연구, 증례보고)
4	· 전문가의견

출처: 김수영 등 2015[14]

13) 서경화, 이선희, 신인순, 임선미, 장지은, 정유민, 박윤형. 근거 및 권고수준 등급화 체계 현황 분석. J Korean Med Assoc 2011;54(7):758-768.

14) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 88p

표 11. SIGN에서 제시한 권고등급 체계

권고등급 단계	구체적 내용
Grade A	· 최소 하나 이상의 메타분석, 체계적 문헌고찰 또는 1++의 무작위임상시험연구가 있고, 표적 모집단에 직접 적용이 가능한 경우 · 주로 1+의 연구로 구성된 근거이고, 직접 표적모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
Grade 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
Grade C	· 2+의 연구로 수정된 근거이고, 직접 표적모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우 · 2++의 평점을 받은 연구로부터 추정된 근거인 경우
Grade D	· 근거의 수준 3 또는 4에 해당되거나 · 2+의 평점을 받은 연구로부터 추정된 근거인 경우

출처: 김수영 등 2015[15)]

1.1.3. 급여 의사 결정

○ 영국에서의 의약품 및 의료기기 사용에 대한 급여 의사결정은 의료기술평가에 근거한 사정(appraisal) 결과를 바탕으로 작성된 NICE 권고문을 통해 이루어지고 있으며, 의료기술평가 결과에 따라 권고는 다섯 가지 유형으로 결정을 내린다(표 12). NICE에서는 의료기술평가 판정 시 질병으로 인한 부담(인구집단에 영향을 미치는 사망률, 상병률), 자원에 미치는 영향(NHS 또는 기타 공공 부분에 미칠 비용적인 영향), 임상적이고 정책적인 중요성(우선순위에 해당하는 주제인지에 대한 여부), 실행함에 있어 부적절한 변화가 존재할 가능성, 권고 생산을 위한 기한에 영향을 미칠 가능한 요인(사안의 긴급한 정도 등), 권고안이 공중보건과 삶의 질에 영향, 건강 불평등의 감소 등에 가져올 가능성을 고려하여 판정하게 된다[16)].

○ 일반적으로 영국의 의료기술은 시장의 자율(free pricing)에 맡기고 있지만 비용효과 분석결과 QALY당 20,000~30,000 파운드인 경우 제조사와 보건부와의

15) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015.92p

16) 박은철, 이효정, 김주영, 남진영, 최재우, 전성연, 주영준, 현인선, 이남경. 국민 안전과 의료기술 조기 활용을 위한 신의료기술평가 제도개선 방안 연구. 보건복지부, 연세대학교. 2016. 48p.

협상을 권고하고 있으며 희귀의약품 또는 생애 마지막(end-of-life)에서 확실한 혁신적 의약품의 경우 QALY당 30,000 파운드를 초과하더라도 급여를 승인하고 있다.

표 12. NICE 의료기술평가 권고

권고	내용
추천함(Recommended)	위원회는 다음과 같을 때 기술 사용에 대한 판단을 내림 기술이 적어도 기존의 기술과 비교할 때 동등한 임상적, 보건의료체계에 대한 편익이 있음이 충분히 확실함 기술이 상당한 임상적, 보건의료체계 편익이 있음이 충분히 확실함
최대한 적합하게 사용 (Optimised)	의료기술 사용으로 부작용의 발생이 가능할 경우, 특정 대상에서만 비용효과적인 경우 제한된 대상에게 최대한 적합하게 사용할 것으로 권고됨. 이 권고에 따라 사용할 시 임상전문가의 조언에 따라 사용되어야 함
오직 연구에서만 사용 (Only in Research)	해당 의료기술의 근거가 불충분한 경우 또는 재평가를 위한 효율성을 입증할 근거가 부족한 경우 충분한 근거가 모이기 전까지 해당기술을 NHS에서 사용할 것을 제안하지 않으며 잘 설계된 연구의 한 부분으로 사용하도록 하고 있음
추천안함 (Not recommended)	새로운 의료기술의 임상효과에 대한 결과가 불충분하거나 현재 NHS하의 임상현자에서 사용과 비교하여 비용-효과성이 부족하다고 판단되는 경우 추천되지 않음
항암제 평가 결과에 따른 사용 추천* (Recommended for use in the CDF)	암 환자의 항암제 사용이 임상적 이득에 대한 초기 근거는 있으나, 비용-효과성에 대한 입증이 추가적으로 필요할 때, Cancer Drugs Fund 내에서 항암제 사용을 권고함

* 항암제 평가에만 적용; CDF, Cancer Drugs Fund

출처: 박동아 등 2018[17)]

17) 박동아, 설아람, 정유진, 최미영, 윤지은, 김윤정, 이나래, 박지정, 손수경, 한해림. 조건부 선별급여 재평가 연구. 2018. 건강보험심사평가원, 한국보건의료연구원67p

1.2. 캐나다

1) 보건의료체계

○ 캐다나의 의료보장은 조세를 기반으로 한 메디케어의 공적건강보험으로 이루어져 있으며, 각 주정부와 준주정부 지역 주민을 위한 독립된 13개 건강보험 플랜에 의해 운영되고 있다. 캐나다 보건법에서는 메디케어를 통해 필수 의료서비스를 전 국민이 본인부담금 없이 제공받을 수 있도록 규정하고 있어 각 주정부 및 준주정부에서 보장하는 의료서비스는 거의 무상으로 제공되며, 민간 의료보험 메디케어 보장 영역을 확대하거나 신속히 메디케어 서비스를 받게 하는 것을 금지하고 있다[18].

2) 의료기술평가제도

가, HTA 개요

○ Canadian Agency for Drugs and Technologies in Health(이하 CADTH)는 캐나다에서 HTA를 수행하는 가장 큰 기관으로 국가적 요구에 기반한 기술평가를 수행하고 있으며, 제약 또는 장비 및 체계 관련 자문 위원회 2개의 위원회(Advisory Committee on Pharmaceuticals or the Devices and Systems Advisory Committee)를 통해 운영된다. 선정된 기술에 대한 HTA는 CADTH 사내, 외부 공공 또는 사설 연구기관에서 수행된다. 따라서 소수의 기술에 대해서만 국가적 건강관리 체계 하의 공식적 임상적, 비용효과성에 대한 리뷰가 이루어지고 있다.

○ CADTH에서는 체계적 문헌고찰 방법을 통해 평가를 수행하며, 급여결정과 관련된 권고는 하지 않는다. 의료기술평가 결과는 의사결정자가 참고할 수 있도록 근거를 제공하는 것이며, 의료서비스 급여 목록 결정권은 주정부가 가지고 있다. 주정부는 의료기술평가 결과와 함께 선호도, 가용자원, 윤리적 고려 등과 같은 지역환경 특성을 고려하여 급여 결정을 한다.

○ 주정부 및 준주정부 단위의 의료기술평가 기구인 OHTAC(Ontario Health

18) 캐나다 보건부 홈페이지. URL:http://www.hc-sc.gc.ca/hcs-sss/medi-assur/index-eng.php

Technology Adivisory Committee)에서는 객관적 근거자료 검토를 통해 급여 결정 권고안을 온타리오주 보건부에 제시하며, 의료기술의 효과성, 비교효과성에 대한 임상적 근거를 체계적 문헌고찰을 통해 수행하고, 합성된 근거들을 Grading for Recommendations Assessment, Development and Evaluation(이하 GRADE)로 평가하여 질적 수준을 4가지로 분류하고 있다.

① 편익이 위험 및 부담 수준보다 높고, 근거 수준도 높음: 특정한 제한 없이 모든 상황에서 적용 가능

② 편익이 위험 및 부담 수준보다 높으나 근거 수준이 낮음: 높은 근거 수준의 연구가 있을 때 의사결정이 달라질수 있음

③ 편익이 위험 및 부담 수준과 동등하나 근거 수준이 높음: 상황과 사회적 가치에 따라 의사결정이 달라질수 있음

④ 편익이 위험 및 부담 수준과 동등하거나 불확실하며 근거 수준이 낮음: 대체기술의 이용이 적절할 수 있음

○ OHTAC는 효과에 대한 근거가 부족한 경우 급여결정을 돕기 위해 현장 평가를 수행하고 있고 이는 주로 GRADE 평가 결과를 기초로 근거 수준이 낮음, 매우낮음, 효과의 잠재성이 매우 중요하다고 평가된 경우 수행된다. OHTAC에서의 권고안은 다음과 같다.

① 의료기술 이용에 대해 보장함

② 이용을 권장하지 않음

③ 특정 임상적 상황 또는 환자만을 대상으로 접근 가능함

④ 현장평가를 통한 향후 평가가 필요함

나, 의료기술평가 예시

○ 캐나다에서 PET 검사의 평가 예시

- 동 평가 보고서에서는 종양질환에서 PET 검사의 임상적 유용성을 평가하고 현존하는 임상 가이드라인 14편에 대한 근거를 제시하였다. 또한, 근거수준 및 권고등급은 미국 흉부외과 학회(American College of Chest Physicians, 이하 ACCP)에서 제시한 근거체계를 이용하여 이에 대한 내용을 다음과 같이 정리하였다[19].

19) Guyatt G, Gutterman D, Baumann MH et al., Grading strength of recommendations and quality of evidence in clinical guidelines. CHEST. 2006;129:174-81.

- ACCP에서는 근거의 수준과 권고 단계를 동시에 표시하며 강한 권고인지 약한 권고인지에 따라 1,2로 분류하고 근거의 강도가 높은 정도에 따라 ABC단계로 분류하고 있다.

표 13. 종양질환에서 PET 검사 사용 지침에 대한 근거수준 및 권고등급

권고등급 및 근거수준	관련 내용
1A	• 강한 권고 • 높은 근거 강도 • 이득과 위해가 균형을 이루지 못함
1B	• 강한 권고 • 중등도의 근거 강도 • 이득과 위해가 균형을 이루지 못함
1C	• 강한 권고 • 매우 낮은 또는 낮은 근거 강도 • 이득과 위해의 균형을 이루지 못함
2A	• 약한 권고 • 높은 근거 강도 • 이득과 위해의 균형
2B	• 약한 권고 • 중등도의 근거 강도 • 이득과 위해의 균형
2C	• 약한 권고 • 매우 낮은 또는 낮은 근거 강도 • 이득과 위해의 균형 또는 불확실성이 존재함

출처: chestnet 홈페이지[20)]

표 14. 가이드라인 개발을 위한 근거의 수준

근거의 수준	
1	체계적 문헌고찰(카테고리 A1) 또는 카테고리 A2로부터 근거한 최소한의 2개의 독립적인 연구
2	체계적 문헌고찰(카테고리 B1) 또는 적어도 2명이 독립적으로 수행한 카테고리 B2의 연구
3	카테고리 A2, B2, C의 연구 1편
4	전문가 의견(카테고리 D)

20) chestnet 홈페이지. Available from URL: http://www.chestnet.org/education/hsp/grading System.php.

문헌 분류	
A1	일관성 있는 결과를 보이는 카테고리 A2 연구에 대한 체계적 문헌고찰
A2	높은 질의 진단법평가연구(참고표준검사와 비교된 중재검사), 치료법과 관련된 높은 질의 RCT 연구(무작위화, 맹검, 완전한 추적관찰, 기저시점의 특성 동질성, ITT 분석)
B1	일관성 있는 결과를 보이는 카테고리 B2 연구에 대한 체계적 문헌고찰
B2	낮은 질의 진단법평가연구(앞서 언급한 특성이 빠진), 또는 낮은 질의 치료법과 관련된 RCT 연구, 또는 비무작위, 코호트, 환자대조군 연구를 포함한 비교연구
C	비교되지 않은 진단법평가연구, 낮은 질의 치료법과 관련된 RCT 연구, 코호트, 환자 대조군 연구, 단일군 연구
D	임상적 경험, 전문가 집단으로부터 얻은 의견

출처: Bipat 등, 2007[21)]

- ACCP 가이드라인 패널은 이득, 위해, 부담, 잠재적 비용을 고려하여 권고하고 있으며, 치료효과에 대한 근거의 질과 불확실성과 관련된 이득과 위해, 부담 간의 균형(trade-off) 여부의 2가지 측면을 고려하여 권고의 강도를 결정하고 있다. 권고는 2개의 레벨, 즉 약함과 강함으로 분류되며, 이득이나 위해, 부담이 명확하다면 grade 1의 강한 권고를 제시하며, 불확실성이 존재하는 경우 grade 2의 약한 권고를 제시하게 된다(표 16). 이러한 2단계의 권고는 단순하고 명확하며 임상의들은 환자의 가치나 선호도의 중요성을 고려해야 함을 제시하고 있다.
- 또한, 방법론적 근거의 질에 있어서는 일관성 없는 결과, 방법론적 단점, 비직접성, 애매모호함 등의 주요 제한점이 없는 무작위 임상시험 연구 및 관찰연구를 높은 근거의 질로 평가하고, 이러한 주요 제한점이 있는 연구유형의 경우 중등도 미만의 근거의 질로 평가한다.

21) Bipat S, van Leeuwen MS, IJzermans JN, Comans EF, Planting AS, Bossuyt PM, et al. Evidence-base guideline on management of colorectal liver metastases in the Netherlands. Neth J Med. 2007;65(1):5-14.

표 15. ACCP의 권고등급 체계의 기준

기준	설명
1	근거의 질과 권고의 등급을 분리함
2	임상의를 위한 단순화 및 투명성
3	충분한(너무 많지 않은) 카테고리
4	가이드라인 개발자를 위한 명백한 방법론
5	가이드라인 개발자를 위한 단순화
6	권고등급 체계의 일반적인 경향성 일치
7	다양한 결과를 위한 다양한 근거 수준에 대한 명백한 접근

표 16. ACCP의 권고 등급

구분	이득 vs. 위해 및 부담	방법론적 근거의 질	내포된 의미
1A/강한 권고, 높은 근거의 질	이득이 위해 및 부담보다 명백히 큼	주요 제한점이 없는RCT 또는 관찰연구를 능가하는 근거	강한 권고로서 제한없이 대부분의 상황에서 대부분의 환자에서 적용가능함
1B/강한 권고, 중등도의 근거의 질	이득이 위해 및 부담보다 명백히 큼	주요 제한점이 있는 RCT (일관성 없는 결과, 방법론적 단점, 비직접성, 애매모호함) 또는 관찰연구로부터 나온 예외적인 강한 근거	상동
1C/강한 권고, 낮은 근거의 질	이득이 명백히 위험과 부담보다 큼	관찰연구 또는 증례연구	강한 권고이지만, 높은 질의 근거가 가능하다면 바뀔수 있음
2A/약한 권고, 높은 질의 근거	이득과 위해, 부담이 균형을 유지함	주요한 제한점이 없는 RCT 또는 관찰연구로부터 나온 근거	약한 권고, 최상의 방침은 환자 또는 사회적 가치, 환경에 따라 바뀔수 있음
2B/약한 권고, 중등도 질의 근거	이득과 위해, 부담이 균형을 유지함	주요한 제한점이 있는 RCT(일관성 없는 결과, 방법론적 단점, 비직접성, 애매모호함) 또는 관찰연구로부터 나온 예외적인 강한 근거	약한 권고, 최상의 방침은 환자 또는 사회적 가치, 환경에 따라 바뀔수 있음
2C/약한 근거, 낮은 질의 근거	이득, 위해, 부담에 대한 불확실성이 존재	관찰연구 또는 증례연구	매우 약한 권고, 다른 대안도 균등하게 타당할 수 있음

표 17. 권고강도에 미치는 영향 요인들

주요 요인	예시
이득, 위해, 비용, 불편성 등을 측정하는 방법론적 근거의 질	여러 높은 질의 RCT는 천식환자에서 스테로이드 흡입 치료의 이득을 설명하고 있으나 단 한편의 증례연구에서 기흉환자에서 흉막 유착에 대한 유용성(효용)을 실험하였음
치료예방의 중요성	심부정맥혈전증 환자에서 혈전용해요법으로 postphlebitic syndrome을 예방하는 것과 폐색전증으로 인한 사망 예방과 대조
치료효과의 크기	심방세동 환자에서 항응고제 vs 위약(RRR, 86%) 보다 일과성허혈발작(TIA) 환자에서 클로피도그렐 vs 아스피린의 뇌줄증 감소(RRR, 8.7%)가 더 작다
치료효과 측정의 정확성	심방세동 환자에서 아스피린(ASA) 치료 vs 위약군은 TIA 환자에서 뇌줄중 예방을 위한 ASA 치료 보다 신뢰구간이 더 넓다
치료와 관련된 위험	급성관상동맥증후군 환자에서 항응고치료를 위한 클로피도그렐과 ASA는 ASA 단독치료 보다 출혈 위험이 더 크다
치료 부담	ASA에 비해 와파린은 투약 용량에 대한 부담이 더 크다. 와파린은 항응고제 강도 모니터링이 필요하며 식이와 관련하여 비타민 K 섭취에 주의해야 함.
목표 사건의 위험	어떤 수술 환자들의 경우 DVT 및 PE의 수술 후 매우 낮은 위험을 가지나 어떤 환자들의 경우 DVT나 PE 비율이 매우 높아질수 있음
비용	TIA 환자에서 클로피도그렐이 ASA 보다더 비쌈
다양한 가치	대부분의 젊고 건강한 사람들은 그들의 삶을 연장시키는 것에 높은 가치를 둘 것이고 노인이나 영유아의 경우 그들의 가치는 다양할 수 있음

표 18. 이득, 위해, 부담, 비용 측정에서 고려되는 주요 요인의 유형

유형	주요 요인
RCT를 바탕으로한 근거의 질을 떨어뜨리는 요인	– 결과의 비일관성 – 근거의 비직접성 – 드문 근거 – 비뚤림이 큰 가능성을 내포한 RCT의 낮은 질의 계획 및 시행
관찰연구를 바탕으로한 근거의 질을 증가시키는 요인	– 효과크기가 큰 경우 – 모든 타당한 교란변수는 명백한 효과를 감소시킴 – 양반응의 증감률

- 또한, 권고강도에 미치는 영향요인으로서 이득, 위해, 비용 등을 측정하기 위한 방법론적 근거의 질, 치료 예방적 의료결과에 대한 중요성, 치료효과의 크기, 치료효과 측정의 정확도(신뢰구간), 치료와 관련된 위험, 치료 부담, 비용, 다양한 환자의 가치 등을 고려하고 있음을 제시하고 있다.
- ACCP 패널은 권고 문구에 대하여 강한 권고의 경우 "~~~ 권고한다(we recommend....)"라고 명시하고, 약한 권고의 경우 "~~~ 제안한다(we suggest....)"라고 강한 권고보다는 다소 명확하지 않은 문구를 제시한다.

3) 급여 의사결정

○ CADTH의 의료기술평가는 임상적 효과성과 비용효과성에 대한 평가결과를 제공하고, 급여결정과 관련된 권고는 하지 않고 있으며[22], 실제 보험급여 여부 결정은 주별로 이루어진다. 의약품, 의료기기(기술)에 대한 보험급여에 관한 최종 의사결정은 주정부, 병원, 사보험사 등 보험자가 결정한다[23].

1.3. 호주

1) 보건의료체계

○ 호주는 정부 주도적인 의료제도를 수행하며, 6개 주정부와 2개 직할자치주(Territory) 정부가 재정을 확보하여 자체법에 따라 의료서비스를 제공한다. 호주의 의료보장은 공공보험과 민간보험의 이중적 구조로 구성되어 있으며, 공공보험으로 1984년부터 전국민 건강보험제도 메디케어(Medicare)를 시행하고 있으며, 민간보험은 선택사항으로 급여범위와 수준이 매우 다양하다.

22) 박은철, 이효정, 김주영, 남진영, 최재우, 전성연, 주영준, 현인선, 이남경. 국민 안전과 의료기술 조기 활용을 위한 신의료기술평가 제도개선 방안 연구. 보건복지부, 연세대학교. 2016. 92p.
23) 김수경, 신상진, 박정수, 김준호, 백영지, 신채민, 김민정, 양장미. 의료기술평가제도의 국제 비교연구. 한국보건의료연구원. 2014. 82p.

2) 의료기술평가제도

가, HTA 개요

○ 호주의 공식적인 의료기술평가는 1980년대에 시작되었으며, 연방과 주정부, 개별병원, 민간기관 및 민간 의료보험 등 다양한 기관과 위원회에서 수행 중이다. 의료기술평가 대상은 의약품(백신 포함), 진단검사(병리학 포함), 의료기기, 외과적으로 시술되는 인공기관, 의료절차, 의료서비스, 수술적 개입, 공중보건 개입 등 모두가 해당되며, 평가기준은 안전성, 유효성 및 비용효과성이다[24].

○ 호주의 의료기술평가는 한국과 유사하게 공보험의 급여목록 등재와 관련해서는 연방 정부에서 관리되고 있으며, 연방 정부의 의료기술평가 결과의 영향력이 가장 크다. 대표적인 의료기술평가 기구로 Medical Services Advisory Committee(이하 MSAC)이 있으며, 신의료기술 및 조기기술에 대한 급여를 목적으로 해당 의료기술이 수가표에 등재될 수 있는지에 대해 보건복지부 자문 역할을 수행하고 있다.

- 호주의 의료기술평가는 연방보건부(Department of Health and Aging)의 치료제품관리국(Therapeutic Goods Administration, 이하 TGA), 의약품급여자문위원회(Pharmaceutical Benefits Advisory Committee, 이하 PBAC), 의료서비스자문위원회(Medical Services Advisory Committee, 이하 MSAC), 인공의료보철, 기기자문위원회(Prostheses and Devices Committee, 이하 PLAC), 예방접종기술자문그룹(Australian Technical Advisory Group on Immunisation, 이하 ATAGI), 호주 신외과수술기술의 안전성 및 유효성 등록(Australian Safety and Efficacy Register of New Interventional Procedures-Surgical, 이하 ASERNIP-S) 등 여러 기관이 상호 의존적으로 연계되어 있다.
- 먼저 의료기술에 포함된 모든 치료재료에 대해 TGA에서 안전성과 효능을 평가하고, TGA 승인을 받은 치료재료는 호주 치료재료 목록(Australian Register of Therapeutic Goods, 이하 ARTG)에 등록된다. 의약품 및 백신은 PVAC에서 평가를 거친 후, 약제급여사업(Pharmaceutical Benefit Scheme, 이하 PBS)이나 국가예방적종프로그램(National Immunisation Program, NIP)에 등재된다.

24) 김수경, 신상진, 박정수, 김준호, 백영지, 신채민, 김민정, 양장미. 의료기술평가제도의 국제 비교 연구. 한국보건의료연구원. 2014.

- 의료기술이나 기기 등은 MSAC에서 안전성, 임상적 유효성, 비용효과성을 평가받는다. MSAC은 의료서비스급여사업(Medicare Benefit Schedule, 이하 MBS)에 새로운 의료서비스를 등재하는 절차에서 필수적이다. 해당 의료서비스의 안전성, 유효성, 비용효과 및 비용에 대한 근거, 해당 의료서비스에 공공자금 지원 여부 및 수가 등을 평가하여 보건부에 자문하는 역할을 하며, 해당 의료서비스가 임상효과나 비용효과가 불확실하여 특정 기간 동안 잠정적으로 지원해야 하는 경우, 합의된 임상적 프로토콜 하에 정의된 데이터가 수집되어 기간 종료 시 MSAC에서 해당 서비스의 재평가를 통보하는 경우에도 자문한다.
- 이미 MBS 목록에 있는 항목 중 새로운 기기들에 대해서는 PLAC의 평가를 받으며, ARTG에 등록된 보철이나 의료기기를 의료기기 급여목록(Prostheses List)에 등재한다. PLAC는 민간보험자가 새로운 보철물과 현존하는 보철물 등 보조기구에 대해 급여하기 위한 급여수준 설정과 새로운 보철 및 기기 등재에 대해 보건부장관에게 권고하지만, 공공재정에 대한 권고는 하지 않는다. PLAC에서는 임상효능을 평가하며, 민간건강보험관계로부터 추천받음 사람으로 구성된다.
- 2010년부터는 의료서비스급여사업(MBS), 약제급여사업(PBS) 및 의료기기 급여목록에 의해 급여를 받고자 하는 신청을 하나로 통일하는 단일진입점(a single entry point) 제도를 도입하였다. 이를 통해 해당 기술을 평가하기 가장 적절한 위원회를 결정하고, 가장 적절한 평가 절차를 제공하는 등의 노력을 통해 시의적절한 평가와 더불어 융합 혹은 복합술의 평가가 가능하게 하여 의료기술평가 운영의 효율성을 개선하고자 하였다. 또한, 신청자는 서로 다른 의료기술평가에 대해 동시에 신청하는 것도 가능하게 되었다. 따라서, 시판허가를 위해 신청하면서 동시에 급여에 관한 신청도 할 수 있으며, 이를 통해 서로 다른 의료기술평가 사업 간에도 관련 정보가 효율적으로 제공될 수 있도록 하였다[25].

25) 최윤정, 채정미, 조수진, 최연미. 신의료기술평가 제도 성과평가 및 운영개선 방안 연구. 건강보험심사평가원. 2013.

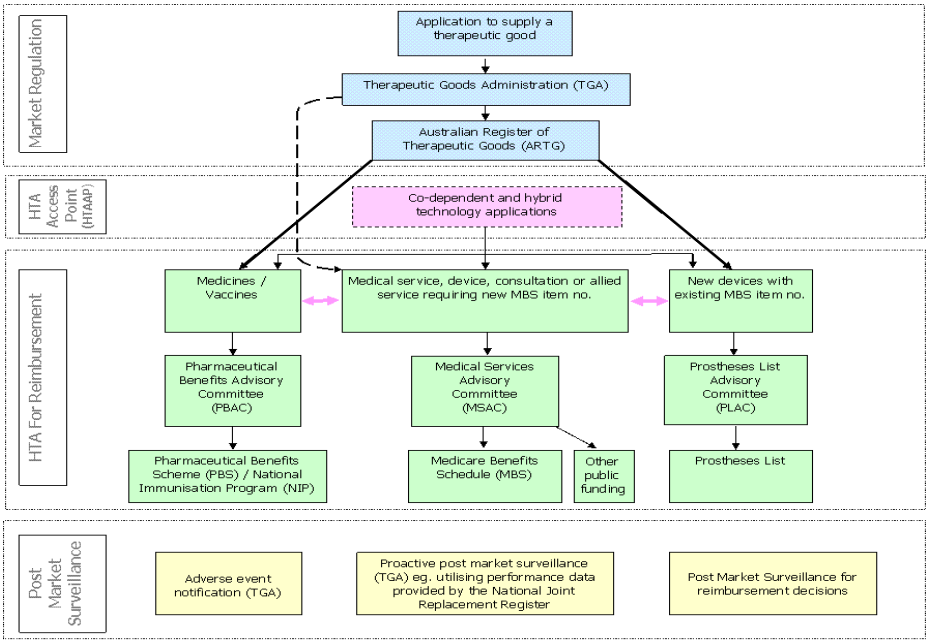

그림 6. 호주 시장 진입 및 급여 결정 시 의료기술평가 절차
(출처: 호주 보건부 홈페이지)

나, HTA 수행 기관 및 역할

1) MSAC (Medical Services Advisory Committee)[26)]

○ MSAC에서는 연방 보건복지부 자문기관으로, 안전성, 유효성, 비용효과성에 대한 근거를 마련하여 급여 결정에 대한 의견을 제시한다.

○ 평가절차

1단계: 의료기술 평가 신청

2단계: 적합성 검토(TGA에서 의료기기 승인을 받았는지 여부 등을 확인)

3단계: PASC 프로토콜 개발

4단계: PASC에 의해 개발된 프로토콜에 근거한 평가

- Contracted Assessment(CA): 실무부서가 평가 질문에 대한 근거를 생성하기 위해 의료기술평가 전문가와 계약을 맺고 평가를 진행함. CA를 위해 MSAC은 4개 기관(University of Sydney, Deakin University, AHTA,

26) 호주 보건부 홈페이지(http://www.health.gov.au/internet/msac/publishing.nsf/Content/about-msac)

ASERNIP-S)과 계약을 맺고 있다.

5단계: 소위원회 평가(제출된 보고서를 검토, 견본과 비교하여 차이여부 및 근거의 불확실성 정도를 직접 급여 범위를 참고하여 정리한다.

6단계: 해당 의료서비스를 권고할 것인지를 판정한다.

7단계: 정부의 결정

2) ASERNIP-S (Australian Safety and Efficacy Register of New Interventional Procedures-Surgical)[27)]

○ 외과의사, 정부, 소비자에게 관련 정보를 제공하기 위해 Royal Australian Colleage of Surgeons (이하 RACS)에 설립된 조직으로, 외과, 심혈관 외과, 산부인과, 안과, 이식, 성형, 재건외과, 비뇨기과, 로봇 수술 등 여러 전문분야의 많은 새 수술기법에 대해 조사하며, 중재적 시술, 기기에 대해 체계적 문헌고찰, 신속검토, Horizon scanning, 요약보고서 등을 담당하는 독립적 기관이다.

다, 근거의 수준 및 권고등급[28)]

○ 진료지침 개발을 위해 Australian Cancer Network(이하 ACN)에서 사용하고 있는 근거의 수준과 National Heart Foundation of Australia(이하 NHFA) 및 Cardiac Society of Australia and New Zealand(이하 CSANZ)에서 제시하고 있는 권고등급은 표19, 표20과 같다. 이 경우 권고등급은 기존 근거에 대한 의견에 차이가 있는 경우도 별도의 등급으로 분류하고 있다. 호주의 NHMRC(The National Health and Medical Research Council)의 경우는 연구설계만으로 근거수준을 결정하며 중재, 진단의 정확성, 예후, 원인, 선별중재에 따라 연구유형을 맞춰서 근거수준체계를 만들어서 사용하고 있다(표 21)[29)].

27) 재인용: 김수경, 신상진, 박정수, 김준호, 백영지, 신재민 등. 의료기술평가제도의 국제 비교 연구. 한국보건의료연구원. 2014

28) 서경화, 이선희, 신인순, 임선미, 장지은, 정유민, 박윤형. 근거 및 권고수준 등급화 체계 현황 분석. J Korean Med Assoc 2011;54(7):758-768.

29) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015.86p

표 19. ACN의 근거수준 체계

근거수준	평가 기준
I	높은 질의 RCT
II	잘 수행된 RCT
III-1	비뚤림 위험이 있는 RCT
III-2	높은 질의 환자대조군/코호트 연구
III-3	잘 수행된 환자대조군/코호트 연구
IV	비분석적 연구(증례연구, 증례보고)

표 20. NHFA/CSANZ의 권고등급 체계

권고등급 단계	구체적 내용
A	임상적용을 권고하는 최상의 근거
B	임상적용을 권고하는 타당한 근거
C	임상적용에 반대하는 타당한 근거
D	불충분한 근거

표 21. 호주 NHMRC 근거수준 체계

수준	중재	진단의 정확성	예후	원인	선별중재
I	'level II' 연구에 대한 체계적 문헌고찰	'level II' 연구에 대한 체계적 문헌고찰	'level II' 연구에 대한 체계적 문헌고찰	'level II' 연구에 대한 체계적 문헌고찰	'level II' 연구에 대한 체계적 문헌고찰
II	무작위 배정 임상시험연구	임상적 상태가 일관된 환자에 대한 독립적, 눈가림법 비교연구	전향적 코호트 연구	전향적 코호트 연구	무작위 배정임상시험연구
III-1	유사 무작위배정 임상시험연구	임상적 상태가 일관되지 않은 환자를 대상으로 한 독립적, 눈가림법 비교연구	All or none	All or none	유사 무작위 배정 임상시험연구
III-2	비교연구 (대조군 있는) · 비무작위실험연구 · 코호트 연구 · 환자 대조군 연구 · 대조군이 있는 간섭 시계열 분석	'level II, III-1'의 기준에 충족하지 못하는 비교연구	무작위배정시험 연구의 단일군에서 예후적 요인 분석	후향적 코호트 연구	비교연구 (대조군 있는) · 비무작위 실험연구 · 코호트 연구 · 환자 대조군 연구
III-3	비교연구 (대조군 없는) · historical control study · 2개 또는 그 이상의 단일군 연구 · 대조군이 없는 개입 시계열 분석	진단적 환자대조군 연구	후향적 코호트 연구	환자 대조군 연구	비교연구(대조군 없는) · historical control study · 2개 또는 그이상의 단일군 연구
IV	실험 전/후 결과의 사례연구	진단분야 연구 (문헌 기준 없음)	사례연구, 서로 다른 stage 환자들에 대한 코호트 연구	단면연구 또는 증례연구	증례연구

출처: 김수영 등 2015[30]

3) 급여 의사 결정

○ 호주 보건부는 MSAC에서 평가한 근거평가 기반 보고를 검토하여 근거정도와 기술의 급여 여부에 관한 권고문을 장관에게 전달하고 우선순위에 따라 급여결정을 시행하고 있다.

○ 해당 근거는 의료기술의 안전성, 효과성, 비용효과성, 총 소요비용과 관련한 근거강도, 공공재원이 해당 의료서비스를 지원해야 하는지 여부, MBS를 통해 재원 조달되는 경우 제안된 MBS 항목에 대한 설명 등의 내용을 포함한다. 이외에 호주 정부는 의료기술이나 중재에 대한 공동체 견해, 영향을 받는 환자그룹, 치료되는 질병의 심각성이나 영향, 효율적인 대안치료 가용성에 대해서도 고려한다[31]. MSAC에서는 의료기술평가 결과에 대한 급여결정과 관련하여 권고안은 세 가지 형태로 제시하고 있으며 이 중 근거가 확실하지 않으나 해당 시술이 비교시술에 비해 안전하고 더 효과적이고 더 비용효과적일 수 있어 등재하는 것이 가능한 경우 MSAC은 해당 기술의 자료수집과 추가적인 평가를 가능하게 하는 임시지원(interim funding)을 권고할 수 있다(표 22).

표 22. MSAC에서 의료기술평가 결과에 대한 권고안

권고안
근거가 확실하여 등재가 가능함
근거가 등재를 지지하지 않음
임시재정(interim funding)을 권고함

출처: 박동아 등 2018[32]

30) 재인용, 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015.87p

31) 김수경, 신상진, 박정수, 김준호, 백영지, 신채민, 김민정, 양장미. 의료기술평가제도의 국제 비교연구. 한국보건의료연구원. 2014.

32) 박동아, 설아람, 정유진, 최미영, 윤지은, 김윤정, 이나래, 박지정, 손수경, 한해림. 조건부 선별급여 재평가 연구. 2018. 건강보험심사평가원, 한국보건의료연구원. 38p

1.4. 미국

1) 보건의료체계

○ 미국의 의료보장은 민간 의료보험과 공공 의료보험의 혼합 운영으로 이루어지고 있으나 민간 의료보험이 지배적이다. 미국의 공공 의료보험은 전 국민 대상이 아닌 노인이나 저소득층 등 특정 범주의 인구만을 대상으로 하며 대표적인 제도로 메디케어(Medicare), 메디케이드(Medicaid), 주정부 어린이건강보험프로그램(State Children's Health Insurance Program, 이하 SCHIP)이 있으며 미국 연방보험청(Center for Medicare and Medicaid Services, 이하 CMS)에서 관장하고 있다.

○ 기타 공적건강보험으로 재향 군인을 위한 의료보장 프로그램, 연방정부 공무원을 대상으로 한 의료보장 프로그램, 인디언을 위한 의료보장 프로그램이 있으며, 민간의료보험은 고용기반 민간보험과 개인 민간보험으로 나뉘며, 대부분이 고용기반 민간보험으로 고용주가 월 보험료 대부분을 지원한다. 대표적인 보험사로 에트나(Aetna), 블루 크로스와 블루 쉴드 플랜(Blue Cross and Blue Shield Plans) 및 보험회사가 네트워크를 형성하여 가입자에게 의료서비스를 제공하는 형태의 건강관리조직(Health Maintenance Organizations, 이하 HMOs)가 있다. 또한, 미국 인구 66%가 민간 의료보험 혜택을 받고 있다.

○ OECD 통계자료에 따르면, 미국의 GDP 대비 지출 의료비는 17.2%로 OECD 회원국 평균 8.9% 보다 높은 수준이며(2017년도 기준), 총 보건의료비 지출 중 개인직접부담 비율이 11.1%로 OECD 평균 대비(20.3%) 낮아 감소추세를 보인다(2016년도 기준)[33].

2) 의료기술평가 제도

가, HTA 개요

○ 미국에서는 병원에서 해당 의료기술의 수용여부를 결정하기 위해 의료기술평가를 수행하고 있으며, 각 보험사는 의료행위에 대해 서로 다른 급여 목록 체계를

33) OECD Health Statistics 2018: Frequently Requested Data(2018, June)

가지고 있음. 따라서 각 병원과 보험사는 의료기술평가조직을 자체적으로 구성하거나 전문 기관에 위탁하여 의료기술평가를 시행하고 있다.

나, HTA 수행 기관 및 프로그램

○ 대표적인 평가기구인 Agency for Healthcare Research and Quality(이하 AHRQ)는 CMS에서 의뢰한 의료기술평가 프로그램, 기관 자체적으로 수행하는 의료기술평가 프로그램(Effective Health Care Program, 이하 EHC)가 있다.

○ HTA 프로그램

- 국가급여결정(National Coverage Determination, 이하 NCD)을 위한 절차는 표 23과 같다.

표 23. HTA 프로그램 시행 절차

단계	구체적 내용
1단계: CMS의 평가 의뢰	NCD 검토가 결정된 주제에 대해 CMS는 필요 시 AHRQ에 의료기술 평가 의뢰
2단계: 연구팀 구성	AHRQ 내부 연구원이 직접 평가 진행 또는 계약된 13개 EPC(Evidence-based Practive Centers) 기관에 위탁하여 평가
3단계: 프로토콜 마련	평가에 포함된 이해관계자가 같이 검토하여 프로토콜 확정
4단계: 분석 및 평가	SR을 통해 평가하며, 프로토콜에 맞게 종합 분석함
5단계: 공개 검토	동료 검토 및 공개 검토를 3주간 진행, 보고서 저자들은 검토 의견에 대한 답변을 3개월 이내에 홈페이지에 공개
6단계: 보고서 공개	검토 내용을 고려한 최종 보고서 홈페이지에 게재

- HTA 프로그램의 근거에 대한 권고사항5가지 권고중 하나로 NCD를 판정하여 결과를 고지한다.
  ① 현재 급여 방법을 그대로 유지함
  ② 급여하지 않음
  ③ 조건없이 급여함: 해당 의료항목 또는 서비스가 모든 메디케어 수혜자에게 합리적이고 필요한 경우
  ④ 특정 조건인 경우에만 급여
    • 특정 상태 환자 또는 인구학적 특성을 지닌 환자에서만 급여

- 특정 기준을 충족하는 의사 또는 요양기관이 제공할 때만 급여
- NCD에서 정한 청구자료 이외 추가자료 수집을 위한 경우에만 급여

⑤ 임상 시험에서 추가적인 자료를 수집하는 것을 조건으로 급여

○ EHC 프로그램

- 제안된 주제에 대해 우선순위에 따라 분석하며, 주제가 선정되면 기존 문헌의 유형과 양에 따라 수행할 연구 종류를 결정하게 된다.

① 비교효과성 검토: 체계적 문헌고찰을 통해 치료방법들의 이이과 위해를 비교함

② 기술요약: 문헌고찰 방법으로 신의료기술에 대한 종합적 검토

③ 연구 보고서: 고찰 연구가 가능할 정도의 문헌적 근거가 부족한 경우 검토

다, 근거수준 및 권고등급[34)]

AHRQ에서 의료기술평가시 별도의 근거수준이나 권고등급을 제시하고 있지는 않으며, 진료지침 개발을 위해는 학회별로 각기 다른 근거수준의 도구를 이용하여 제시하고 있다. 근거의 수준과 함께 이익과 손실의 크기를 고려한 U.S. Preventive Services Task Force(이하 USPSTF)에서 근거수준을 고려시 포함하고 있는 요소와 권고등급 체계는 아래 표 24, 표 25와 같다.

표 24 USPSTF의 근거수준 고려요소

내용
1. 연구들이 핵심질문에 적절한 연구 설계를 가지고 있는가?
2. 현존하는 연구들의 질이 어느 수준까지 높은가?(내적 타당도의 수준)
3. 연구결과들은 미국 일차진료 대상집단 및 상황에 어느 정도 일반화시킬수 있는가?(외적 타당도의 수준)
4. 핵심질문과 관련된 연구들의 수
5. 연구결과들의 일관성
6. 결론을 내리는데 도움이 되는 추가적인 요소들을 제공하는가?(양-반응효과의 존재, 생물학적 모형에 맞춤)

출처: 김수영 등 2015[35)]

34) 서경화, 이선희, 신인순, 임선미, 장지은, 정유민, 박윤형. 근거 및 권고수준 등급화 체계 현황 분석. J Korean Med Assoc 2011;54(7):758-768.

35) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 88p

표 25. USPSTF의 권고등급 체계

권고등급 단계	구체적 내용
A	해당 서비스를 권고한다. 순이득(net benefit)에 대한 근거가 충분하다.
B	해당 서비스를 권고한다. 순이득에 대한 근거는 중등도 또는 중등도와 충분함 사이에 있다.
C	해당 서비스를 선택적으로 제공하거나 또는 전문가 판단과 환자 선호도에 따른 특정 개인에 시행할 것을 권고한다. 순이득에 대한 근거가 작다는데 대한 중등도의 확신이 있다.
D	해당 서비스를 권고하지 않는다. 순이득보다 위해가 크다는 데 대한 중등도 또는 높은 수준의 확신이 있다.
I statement	이득과 위해의 저울질을 평가하기 위한 현재의 근거가 불충분하다. 근거는 부족하고 낮은 질을 가지거나 상충하고, 이득과 위해의 균형을 확인할 수 없다.

출처: 김수영 등 2015[36)]

3) 급여 의사 결정

○ 미국에서의 급여결정은 CMS에서 담당하고 공식 요청을 접수한 후 검토를 거쳐 국가급여결정을 발표한다. 급여결정 결과는 표 26과 같으며, 이 중 특정 조건인 경우에만 급여하거나 임상연구에서 추가적인 데이터를 수집하는 조건으로 급여하는 경우가 Coverage with Evidence Development (CED)에 해당된다[37)]. 급여결정 기준에는 그 실험이 실험적인지, 일반적으로 용인되는지, 상대적으로 안전한지, 비용이 얼마나 드는지, 비용이 적게 들지만 같은 편익을 나타내는 다른 시술이 있는지, 접근 가능한지, 임상적으로 효과적인지를 포함하고 있다[38)].

36) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 93p

37) 박동아, 설아람, 정유진, 최미영, 윤지은, 김윤정, 이나래, 박지정, 손수경, 한해림. 조건부 선별급여 재평가 연구. 건강보험심사평가원, 한국보건의료연구원. 2018. 12p.

38) 박실비아, 박은자, 채은미, 이예슬. 신의료기술의 도입과 확산에 관한 정책분석. 한국보건사회연구원. 2014. 94p.

표 26. 미국의 급여결정 결과

급여결정 결과	설명
현재 급여 방법을 그대로 유지	현재의 지역 또는 국가 급여 사항에 변동이 발생하지 않음
급여하지 않음	해당 의료 항목 또는 서비스가 합리적이며 필요하다고 결론을 내리기에는 의학적 근거가 적합하지 않아 메디케어 수혜자에 대한 급여가 허용되지 않음
조건없이 급여함	해당 의료 항목 또는 서비스가 모든 메디케어 수혜자에 대해 합리적이고 필요하다고 결론을 내리기에 적절함
특정 조건인 경우에만 급여함	해당 의료 항목 또는 서비스가 아래와 같은 상황에 대해서만 합리적이고 필요하다고 결론을 내리기에 적절함 ① 특정 상태의 환자 또는 인구학적 특징을 지닌 환자에게만 급여 ② 특정 기준을 충족하는 의사 또는 요양기관이 제공할 때에만 급여 ③ NCD에서 정한 청구자료 이외 추가 자료 수집을 위한 경우에만 급여
임상연구에서 추가적인 데이터를 수집하는 것을 조건으로 급여	의료기술의 추가 자료의 수집을 위해 사전에 규정된 절차에 따라 제공될 때에만 급여함

출처: 박은철 등 2016[39)]

39) 박은철, 이효정, 김주영, 남진영, 최재우, 전성연, 주영준, 현인선, 이남경. 국민 안전과 의료기술 조기활용을 위한 신의료기술평가 제도개선 방안 연구. 보건복지부. 연세대학교. 2016. 71-720p

1.5. 국가별 제도 비교

주요 4개국에 대한 의료기술평가 제도 비교 내용은 표 27 및 표 28과 같다.

표 27 주요국 의료기술평가 결과 비교

구분	영국	캐나다	미국	호주
HTA 운영	지역별로 다르게 운영	주단위로 지역적 상황을 고려한 의료기술평가 수행	각병원과 보험사에서 자체적으로 구성하거나 전문기관에 위탁	연방과 주정부, 개별병원, 민간기관 및 민간의료보험등 다양한 기관과 위원회에서 수행
대표 HTA 기관	NICE	온타리오주 (OHTAC)	AHRQ	MSAC
HTA 결과에서 근거수준, 권고등급 제시 관련	NICE에서 제시한 보고서에서는 별도의 근거수준, 권고등급을 제시하지 않음.	GRADE 이용	AHRQ에서 제시한 보고서에서는 별도의 근거수준, 권고등급 제시하지 않음.	NHMRC에서 제시한 근거수준체계로 제시
의료기술평가 결과 권고	추천함/최대한 적합하게 사용/오직 연구에서만 사용/추천안함	의료기술 이용 보장/이용을 권장하지 않음/특정 임상적 상황 또는 환자만을 대상으로 접근 가능/현장평가를 통한 향후 평가 필요	현재 급여방법 유지/급여하지 않음/조건없이 급여/특정 조건인 경우에만 급여/임상연구에서 추가적인 데이터를 수집하는 것을 조건으로 급여	급여가 확실하여 등재가능/근거가 등재를 지지하지 않음/임시재정을 권고

표 28. 주요국의 의료기술평가 권고결정 고려요인

구분	영국	캐나다	미국	호주
의사결정시 고려요인	근거의 질 -GRADE사용	종합적인 임상적 이득 - Effectiveness - Safety - Burden of illiness - Need	근거수준 - 높음/적정/낮음	근거수준
	이득과 위해 trade-off	Consistency with expected societal and ethical values	총 편익(benefit) 규모	일관성
	Equity	Value for money	-	임상적 효과
	Economic and resource use	Feasibility of adoption into the health system - economic feasibility - Organizational feasibility	-	일반화 가능성
	Availability	-	-	적용가능성

1.6. 기관, 임상학회별 근거수준 및 권고등급 체계조사

기관별, 임상학회별로 사용하고 있는 근거수준 및 권고수준의 등급화 체계는 다음과 같다[40].

먼저 근거수준은 연구설계방법, 연구의 질, 일관성, 한계점 및 연구설계 방법의 강도에 따라 4가지 유형으로 나누어 볼 수 있다. 유형 1의 경우 연구설계방법 뿐만 아니라 전반적인 연구의 질적인 측면 및 일부 기관들에서는 일관성에 관한 기준도 적용하여 분류하고 있었고 SIGN, NICE, Singapore Ministry of Health(SMOH) 등의 23개 기관에서 활용하고 있었다. 유형 2의 경우는 연구의 설계방법만을 등급 분류기준으로 고려하고 있었으며 비무작위할당 통제실험 및 준-실험연구로부터 얻은 근거의 등급도 포함하고 있었고, British Association of Sexual Health and HIV(BASHH), Royal College of Obstetricians and Gynaecologists (RCOG), Registered Nurses' Association of Ontario(RNAO) 등의 12개 기관에서 활용하고 있었다. 유형 3의 경우는 연구설계방법과 연구의 한계점을 분류기준으로 고려하고 있었으며 American Academy of Pediatrics(AAP), Global Initiative for Chronic Obstructive Lung Disease(GOLD)등의 7개 기관에서 활용하고 있었다. 유형 4의 경우 연구설계방법의 강도를 분류 기준으로 고려하고 있었으며, American Dietetic Association (ADA), American College of Radiology(ACR)의 2개 기관에서 활용하고 있었다(표 29).

권고 수준은 연구의 질, 일관성, 적용가능성, 근거의 질, 권고수준의 강도, 이익과 손실의 균형 등의 기준에 따라 5가지 유형으로 구분될 수 있다. 유형 1의 경우 연구의 질 및 근거수준을 기준으로 등급을 구분하고 있으며 일부 기관에서는 일관성 및 적용가능성을 등급분류기준으로 고려하여 SIGN, SMOH, NICE 등의 9개 기관에서 개발한 진료지침에서 활용하고 있다. 유형 2의 경우 근거의 수준과 권고에 대한 강도의 크기를 감안하여 등급을 구분하고 있으며, American Society for Gastrointestinal Endoscopy(ASGE), ADA, Heart Failure Society of American(HFSA), American Psychiatric Association(APA)를 포함한 15개의 진료지침에서 활용하고 있다. 유형 3의 경우 근거수준에 대한 내용을 기준으로 구분하고 있으며, American Academy of Orthopaedic Surgeons(AAOS), Society of Obstetricians and Gynaecologists of Canada(SOGC) 등의 11개 진료지침에서 활용

40) 서경화 등. 근거 및 권고수준 등급화 체계 현황분석. J Korean Med Assc 2011. 54(7): 758-768.

하고 있다. 유형 4의 경우 근거의 수준과 함께 이익과 손실에 대한 크기의 비교를 기준으로 등급을 구분하고 있으며 U.S.Preventive Service Task Force(USPSTF), American Gastroenterological Associaton Institute(AGA), European Society of Cardiolog(ESC), Committee on Perioperative Evaluation/Brazilian Society of Cardiology(CAPO/BSC) 4개 기관에서 진료지침 개발시 적용하고 있다(표 30).

표 29. 근거의 수준 유형

구분	내용
유형 1	High quality RCT, Well conducted RCT, RCT with high risk of bias, High quality case-control/cohort study, Low quality RCT or cohort study, Outcome research(ecological studies), Systematic review of case-control study, Well conducted case-control/cohort study, Case-control/cohort study with high risk of bias, Non-analytic studies(case series/case report), Expert opinion
유형 2	Meta-analysis of RCT, At least on RCT, At least one controleed study without randomisation, Quasi=experimental study, Non-experimental descriptive study(case-control study), Expert committee reports/opinion
유형 3	Well-designed RCT, RCT with minor limitation(inconsistent results/methodological flaws), Overwhelming evdence from observational study/RCT with serious flaws, Observational study(case-control/cohort study), Expert opinion/case reports, Exceptional situations(validating studies cannot be performed)
유형 4	Studies of strong design, Studies of strong design with minor methodological concersn, Studies of weak design, No reliable evidence/expert opinion, No evidence

표 30. 권고의 수준 유형

구분	내용
유형 1	· Evidence from level 1++/1+, directly applicable, consistency results · Evidence from level 2++, directly applicable, consistency results · Evidence from level 2+, directly applicable, consistency results · Evidence level 3 or 4 · Good practice points based on clinical experience · No conclusive level of evidence category

구분	내용
유형 2	· Strong recommended, good evidence · Strong recommended, can apply to most clinical settings · Strong recommended, fair evidence · Intermediate-strength, may change in stronger evidence · Intermediate-strength, may differ depending on circumstances · Weak recommended, alternative approaches in better circumstances
유형 3	· Very weak recommended, likely alternative approaches in better circumstances · Weak recommended, likely to change aas data become available · No recommended · Recommend against routine/prevention setting · Based on Expert opinion
유형 4	· Good evidence to recommend the clinical action · Fair evidence to recommend the clinical action · Existing evidence is conflicting · Fair evidence to recommend against clinical action · Good evidence to recommend against clinical action · Inefficient evidence
유형 5	· Net benefit is substantial · Net benefit is moderate · Net benefit is small · No net benefit or harms outweigh the benefit · Insufficient evidence

2. GRADE 체계

2.1. 개요

Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working 그룹에서는 의학 연구의 질 이외의 결과의 신뢰성에 영향을 미치는 관점까지 고려한 권고의 강도 체계를 만들었다(강현, 2016). 근거수준과 권고강도를 등급화한 GRADE는 체계적 문헌고찰부터 보건의료 가이드라인 및 권고안을 개발하고 실행하는 투명하고 구조화된 과정을 제공한다. 결과적으로 GRADE는 임상의와 환자에

게 임상치료에 대한 지침을 제공하고 정책결정자들에게는 의사결정에 필요한 근거를 제공할 수 있다(최윤정 등, 2014; Guyatt et al., 2011). 최근까지, WHO, 코크란, NICE를 비롯하여 전 세계적으로 100여개 이상의 전문 기관이 근거에 기반한 권고를 결정하는데 GRADE 방법론을 활용하고 있다(Alonso-Coello et al., 2016; Dahm et al., 2017).

GRADE의 기본적인 과정은 수집된 연구들에서 결과지표별로 근거의 질을 평가하여 근거수준을 평가하는 단계와 근거수준을 바탕으로 이득과 위해의 균형(balancing), 가치와 선호도 등을 고려하여 권고의 방향과 강도를 정하는 권고등급 결정 단계로 이루어져 있다.

GRADE에서는 근거의 확실성을 평가하고, 근거로부터 권고 또는 의사결정을 잇는 과정을 보다 투명하고 객관성있게 제시하기 위해 근거에 기반한 의사결정 틀(evidence-to-decision (EtD) framework)을 개발하여 활용하고 있다(Dahm et al., 2017).

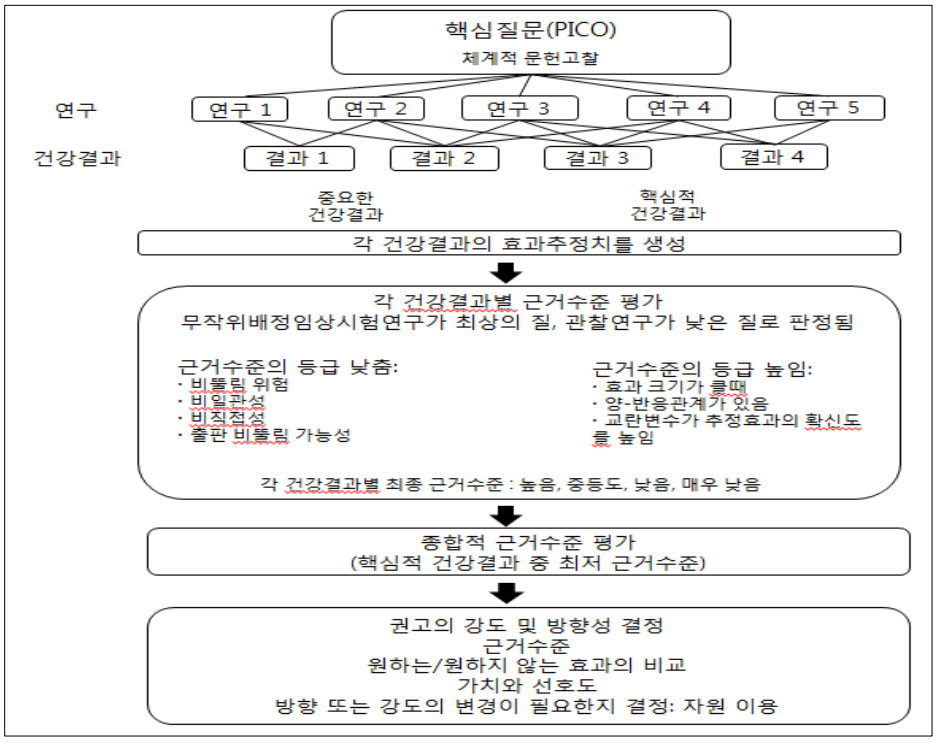

그림 7. 권고안 도출 관련 GRADE 과정의 도식화

(재인용: 한국보건의료연구원 2015, 원저: Guyatt et al 2011)

2.2. 근거수준 평가

근거수준은 현재까지 근거로 해당 중재의 효과에 대해 확신하는 정도를 말한다. 근거수준을 결정할 때 고려되는 요소로는 ① 연구설계, ② 문헌의 질, ③ 근거의 양, ④ 근거의 일관성, ⑤ 근거의 직접성 등이 있으며, GRADE에서는 모든 요소를 사용하여 근거수준을 판단한다. 이 체계에서는 각각의 연구 결과가 아니라 근거 전체에서 관찰된 효과가 실제 효과에 가까울 것으로 판단되는 정도에 따라 “높음”, “중등도”, “낮음”, “매우 낮음”의 4개의 범주로 등급을 부여하여 근거의 질을 평가한다(강현, 2016).

GRADE에서 근거수준은 연구설계에 따라 우선적으로 결정된다. 무작위배정 비교임상시험의 경우 ‘높음(high)’, 관찰연구인 경우 ‘낮음(low)’으로 배정한다. 이후 연구설계에 따른 비뚤림의 위험(risk of bias)이 있거나 근거가 직접성이 없을 때(indirectness), 연구의 정확성이 떨어질 때(imprecision), 설명되지 않은 이질성이 있을 때(inconsistency), 출판 비뚤림의 위험성(publication bias)이 있을 때 등급을 1 또는 2단계 낮춘다. 반면, 효과크기가 크거나(large effect), 교란변수의 가능성이 적을 때(palusible confounding), 용량-반응(dose response)을 보일 때 등급을 올려 근거수준의 등급을 조절한다(강현, 2016; 임상진료지침 개발 매뉴얼, 2011).

표 31. GRADE에서 근거수준 의미

근거수준	정의
높음 (High) ⊕⊕⊕⊕	추후 연구로 효과추정치에 대한 확신정도가 바뀔 가능성은 매우 낮다
중등도 (Moderate) ⊕⊕⊕O	추후 연구로 효과추정치의 확신 정도에 중요한 영향을 미칠 것이며 추정치가 바뀔수도 있다
낮음 (Low) ⊕⊕OO	추후 연구는 효과추정치의 확신정도에 중요한 영향을 미칠 가능성이 매우 크며, 추정치가 바뀔 가능성이 높다
매우 낮음 (Very Low) ⊕OOO	효과에 대한 어떤 추정도 불확실하다

출처 : Balshem et al 2011

표 32 GRADE에서 근거수준 평가기준

연구설계	초기 근거수준	등급 낮춤	등급 높임	근거수준
무작위연구	높음 →	비뚤림 위험 심각: −1 매우 심각:−2 비일관성 심각: −1 매우 심각: −2 비직접성 심각: −1 매우 심각: −2 비정밀성 심각: −1 매우 심각:−2 출판비뚤림 강하게 의심됨:−1	효과크기 큼: +1 매우 큼:+2 양 반응관계 있음: +1 교란변수 효과 추정의 확신도를 높임: +1	높음 ⊕⊕⊕⊕ 중등도 ⊕⊕⊕○ 낮음 ⊕⊕○○ 매우낮음 ⊕○○○
관찰연구	낮음 →			

재인용: 한국보건의료연구원, 2015[41], 원저: Guyatt et al., 2011[42]

2.3. 권고의 도출

GRADE의 권고는 방향성(권고함, 권고하지 않음)과 강도(강함, 약함)을 토대로 크게 4가지 범주(강하게 권고하지 않음, 약하게 권고하지 않음, 약하게 권고, 강하게 권고)로 구분될 수 있으며(표 33), 권고등급은 원하는 효과와 원하지 않는 효과의 비교, 효과의 확신도(근거수준), 가치 및 선호의 신뢰도, 자원이용을 통해 결정될 수 있다(표 34)[43].

41) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 230p.

42) Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables, J clin Epidemiol. 2011:64(4):383-94.

43) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 310-13p.

표 33 GRADE의 권고 강도 4가지

GRADE의 권고 강도	내용
강한 권고 (strong recommendation)	· 중재의 원하는 효과가 원하지 않는 효과보다 훨씬 크거나, 또는 반대로 원하지 않는 효과가 원하는 효과보다 훨씬 큼 · 주의할 점은 강한 권고가 반드시 높은 우선순위를 가지는 권고는 아님 · 해당 중재에 적용을 받는 대부분 혹은 모든 사람들에게 최선으로 적용할 수 있다는 의미임
약한 권고 (weak recommendation)	· 중재의 원하는 효과가 원하지 않는 효과보다 약간 큰 경우 또는 반대로 원하지 않는 효과가 원하는 효과보다 약간 큰 경우를 의미함 · 이는 모든 사람들에게 해당 중재를 최선으로 적용할 수 있다는 의미는 아님 · 다른 유사표현은 다음과 같음 – 조건부(conditional, 환자 선호도, 가용자원이나 환경에 의해) – 재량에 의해(discretionary, 환자 또는 의사의 의견) – 충분할 때(qualified, 결정을 내릴 수 있는 상황에 대한 설명에 의해)
연구에서만 권고 (recommendation only in research)	· 유망하거나 혹은 신개발 중재의 경우 위해 또는 근거가 불충분한 경우가 있음. 이 표현은 다음의 세가지 조건을 모두 충족할 때 적용됨 ① 해당 중재를 지지하거나 반대할 수 있는 근거가 매우 불충분할 때 ② 추가 연구를 통해서 해당 중재 효과의 불확실성을 상당 부분 줄일 수 있는 가능성이 있을 때 ③ 추가 연구가 연구에 예상되는 비용에 해당하는 가치가 있다고 생각될 때
권고하지 않음 (recommendation)	권고를 내리지 못하는 다음의 3가지 이유 중 한 가지에 해당되고, 연구에서만 권고하는 것이 부적절할 때 적용할 수 있음. ① 효과 추정치의 신뢰도가 너무 낮아서 권고의 추측성이 너무 클 때 ② 효과 추정치의 신뢰도와 관계없이, 효과크기의 저울질에서 차이가 거의 없으며, 가치와 선호도 및 자원이용이 알려져 있지 않거나 다양성이 너무 커서 권고의 방향성을 결정하는 데에 큰 어려움이 있을 때 ③ 2가지 선택 대안이 너무 다른 원하지 않는 결과를 가져오고, 개별 환자의 해당 결과에 대한 반응 역시 너무 달라서 일반적인 가치와 선호도 관점에서 결론을 내릴 수 없을 때

재인용: 한국보건의료연구원, 2015[44)]

44) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 310-12p.

표 34. 권고등급을 결정하는 요인

요인	설명
원하는/원하지 않는 효과의 비교	원하는 효과와 원하지 않는 효과의 차이가 클수록 강한 권고를 내릴 가능성이 높아짐 그 차이가 작거나 편익에 대한 확신이 낮을수록 약한 권고를 내릴 가능성이 높아짐
효과의 확신도 (근거수준)	근거수준이 높을수록 강한 권고를 내릴 가능성이 높아짐
가치 및 선호의 신뢰도	가치와 선호의 가변성이 클수록, 또는 가치와 선호에 대한 신뢰도가 낮을수록 약한 권고를 내릴 가능성이 높아짐
자원 이용	중재법의 비용이 클수록(자원 이용이 많을수록) 강한 권고가 내려질 가능성은 낮아짐

재인용: 한국보건의료연구원, 2015[45)]

2.4. 근거에 기반한 의사결정 틀

근거에 기반한 의사결정(EtD) 틀의 개발 목적은 임상적 권고, 보험급여(coverage) 결정 및 보건시스템 또는 공중보건 권고/결정 맥락 안에서 의사결정을 내릴 때 구조화되고 투명한 방법으로 근거를 사용하도록 돕는 것이다(Alonso-Coello et al., 2016). EtD 틀(framework)은 질문 구조화, 근거의 평가 및 결론을 도출하는 세가지 단계의 공통된 구조를 가지고 있지만, 의사결정의 각 유형별로 framework에는 약간의 차이점이 있다.

표 35. 의사결정 유형별 EtD frameworks 기준

구분	임상적 권고 –inidvidual 관점	임상적 권고 – population 관점	급여 결정	보건체계 및 공중보건에 대한 권고/결정	진단, 선별, 기타 검사*
문제 우선순위	문제가 우선사항인가?				
검사 정확성	적용안됨				검사는 얼마나 정확한가?

45) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 313p.

<table>
<tr><th>구분</th><th>임상적 권고 –inidvidual 관점</th><th>임상적 권고 – population 관점</th><th>급여 결정</th><th>보건체계 및 공중보건에 대한 권고/결정</th><th>진단, 선별, 기타 검사*</th></tr>
<tr><td rowspan="2">이득과 위해</td><td colspan="5">예상되는 긍정 효과(desirable effects)는 얼마나 큰가?</td></tr>
<tr><td colspan="5">예상되는 부정 효과(undesirable effects)는 얼마나 작은가?</td></tr>
<tr><td>근거의 확실성 (근거수준)</td><td colspan="4">효과에 대한 근거의 확실성이 전반적으로 어떠한가?</td><td>근거의 확실성:
– 검사 정확도?
– 검사에 있어 중대/중요한 직접적 이득, 부작용 효과, 부담이 있는가?
– 검사 결과에 따른 관리 효과가 있는가?
– 검사 결과와 관리 의사 결정간의 연계가 있는가?
– 검사의 효과?</td></tr>
<tr><td>가치</td><td colspan="4">주요 의료결과를 얼마나 중요하게 여기는지, 즉 가치에 대해 중요한 불확실성 또는 가변성이 있는가?</td><td>검사의 부작용, 부담(burden), 검사결과에 따른 임상적 관리 결과 등을 포함하여 주요의료결과의 가치에 중요한 불확실성 또는 가변성이 있는가?</td></tr>
<tr><td>효과간 저울질(balancing)</td><td colspan="4">긍정 효과와 부정 효과를 고려할 때, 중재법 또는 비교법 중 어느 쪽이 더 선호되는가?</td><td>긍정효과와 부정효과를 고려할 때, 중재검사 또는 비교검사 중 어느쪽이 선호되는가?</td></tr>
<tr><td>자원사용</td><td>–</td><td colspan="4">필요자원(비용)은 얼마나 큰가?</td></tr>
<tr><td>–</td><td>필요자원(비용)에 대한 근거의 확실성은 어떠한가?</td><td colspan="4">자원의 요구사항(비용)에 대한 근거의 확실성은 어떠한가?</td></tr>
<tr><td>중재의 비용효과(순 이득 대비 본인부담비용)가 중재 또는 비교에 유리한가?</td><td>중재의 비용효과성이 중재 또는 비교에 유리한가?</td><td colspan="4">중재의 비용효과성이 중재 또는 비교에 유리한가?</td></tr>
</table>

구분	임상적 권고 –inidvidual 관점	임상적 권고 – population 관점	급여 결정	보건체계 및 공중보건에 대한 권고/결정	진단, 선별, 기타 검사*
형평성	–	건강형평성에 미치는 영향은 어떠한가?			
수용성	중재가 환자, 보호자, 보건의료제공자에게 수용될 수 있는가?	중재가 주요 이해관계자에 수용될 수 있는가?			
실행가능성	중재가 환자, 보호자, 보건의료제공자에게 실행가능한가?	중재가 실행가능한가?			

* 검사는 개인/집단적 관점에서 임상 및 공중보건 권고사항을 포함함

출처 : Alonso-Coello et al 2016

가, 질문 구조화

GRADE에서 근거로부터 의사결정을 내리기 위한 첫 번째 단계는, 의사결정을 수행하고자 하는 주제에 대한 질문을 PICO(사용대상/ 중재/ 비교/ 주요 의료결과)의 형식에 맞게 상세하게 설정한다. 또한, 권고 또는 의사결정이 왜 필요한지 보다 이해할 수 있도록 질문의 핵심 배경 정보, 의사결정자의 관점(perspective), 사용대상의 하위그룹에 대해서도 명확하게 명시한다.

나, 근거의 평가

근거의 평가는 EtD 틀 평가기준(표 35)을 이용하여, 각 기준에 대한 판단, 연구근거, 추가적인 고려사항에 대해 평가한다. 의사결정의 유형에 따라 기준의 조작화(operationalisation) 차이는 있긴 하지만, 대부분의 평가기준은 표 35에서와 같이 유사하다. 평가기준은 ① 문제가 우선순위인지, ② 원하는 효과와 원하지 않는 효과의 크기, ③ 근거의 확실성, ④ 주요 결과에 대한 환자(또는 간병인과 같이 영향받는 사람) 측면에서의 가치(의료결과의 중요성), ⑤ 원하는 효과와 원하지 않는 효과간의 저울질(balancing), ⑥ 자원이용, ⑦ 수용가능성 및 ⑧ 실행가능성을 포함하고 있다. 이밖에도 인구집단의 관점에서 의사결정을 수행하는 경우에는 형평성

에 미치는 영향을 포함한다.

검사기술에 대한 근거 평가시에는, RCT 또는 관찰연구에서 중요 의료결과에 대해 대체검사로서의 영향에 대한 근거가 없는 경우 추가적인 기준이 요구된다. 즉, 검사의 정확성, 검사의 효과(비침습적 검사에 따른 부작용 감소 효과와 같은 직접적인 효과, 검사결과에 따른 치료/관리결정으로 인한 결과와 같은 간접적인 효과)에 대한 근거의 확실성을 고려사항으로 포함한다(Alonso-Coello et al., 2016).

연구 근거는 체계적 문헌고찰로부터 파생된 연구결과를 토대로 제시하며, 연구 근거 이외에 수집된 데이터, 가정(assumptions), 판단에 사용된 논리와 같은 다른 근거를 추가적인 고려사항으로 평가한다.

다, 결론도출

모든 기준에 대해 내려진 평가를 토대로, 의사결정자는 권고의 강도(강함 vs 약함) 또는 의사결정의 유형에 대한 결론을 도출한다. 또한, 의사결정자는 권고 또는 의사결정 내용을 간결하고 명확하게, 실행가능한 방식으로 기술하여 정당성을 제시한다. 권고 및 의사결정의 유형은 다양하다. 임상적 권고시에는 강한 권고 또는 약한 권고가 적합한 형태의 결론 도출 방법이라고 볼 수 있지만, 보험급여, 보건체계 또는 공중보건 의사결정시에는 적절하지는 않다. 보험급여결정에서는 급여 미적용(no coverage), 근거 생산 조건부 급여(coverage with evidence development(in the context of research)), 가격협상을 통한 급여(coverage with price negotiation), 제한적 급여(restricted coverage), 급여(full coverage)로 결론 도출이 가능하다.

표 36. Evidence to Decision(EtD) frameworks 기준에 대한 질문 및 세부지침

기준	질문	질문 설명	세부지침
1. 문제	문제의 우선순위는?	해당 중재법과 관련된 문제가 심각하고(또는 비용이 많이 드는), 긴급한가? 아니면 정치적 또는 정책적 지원이 필요한 문제인가?	문제가 심각할수록, 이러한 문제에 대처하는 중재법을 우선시 할 가능성은 높아진다(예컨대 치명적이거나, 장애를 유발하는 질병은 경미한 질병보다 우선시 될 가능성이 높다) 많은 사람이 영향을 받는 문제일수록, 이렇나 문제에 대처하는 방안을 우선시 할 가능성은 높아진다.
2. 가치	주요 결과에 대한	영향을 받는 사람들은 각각	가치 판단의 차이로 서로 다른 결정

기준	질문	질문 설명	세부지침
	가치판단에 중요한 불확실성이나 가변성이 존재하는가?	주요 결과를 얼마나 중요하게 생각하는가? 각각의 주요 결과를 어떻게 평가하는지에 대한 불확실성, 또는 결과에 대한 사람들의 다양한 관점(가변성이 의사결정을 좌우할 만큼 큰가?	을 내릴 가능성이 높을수록, 우선적인 방안으로 합의될 가능성은 낮아진다. (또는 해당 중재법에 영향을 받는 사람들의 가치에 대한 근거를 획득하느 ㄴ일이 더욱 중요해진다.) 여기서 가치(value)라는 것은 관심결과의 상대적인 중요도(사람들이 이러한 각각의 결과를 중요시하는 정도)를 의미한다. 이러한 가치를 때때로 '효용 가치' (utility values)라고 일컫는다.
3. 효과의 확신도 (근거수준)	효과와 관련된 근거의 확실성은 전반적으로 어떠한가?	효과와 관련 근거수준은 어떠한가? 연구를 통해 확인된 근거는 모든 핵심 겨로가에 얼마나 긍정적인 영향을 미칠 수 있겠는가? 예컨대, 의사결정에 영향을 미칠 수 있는 연구 결과(근거)는 실제 효과와 얼마나 차이가 있는가?	권고를 이끄는 핵심 결과에 대한 근거가 불확실할수록, 해당 중재법이 권고될 가능성은 낮아진다.(만약 권고된다면, 예비연구 또는 영향력 평가의 중요성이 커질 수 있다.)
4. 원하는(긍정) 효과	원하는 효과의 크기는 어느 정도인가?	해당 중재법의 원하는 효과(건강을 비롯한 여러 가지 편익 포함)는 얼마나 강력한가? 그리고 이러한 효과의 크기는 어느 정도인가?(얼마나 많은 사람들이 편익을 취할 것인가? 또는 개선효과가 얼마나 클 것인가?)	편익이 클수록, 해당중재법이 권고될 가능성은 높아진다.
5. 원하지 않는(부정)효과	원하지 않는효과의 크기는 어느 정도인가?	중재법의 원하지 않는 효과는 얼마나 강력한가? 그리고 이러한 효과의 크기는 어느 정도인가?	위해가 클수록, 해당 중재법이 권고될 가능성은 낮아진다.
6. 효과의 비교	원하는 효과와 원하지 않는 효과를 모두 고려했을 때, 해당 중재법이 비교 중재법보다 더 선호되는가?	주요 결과에 대한 개인의 가치 평가, 실질적인 효과, 효과 추정치의 확실성, 할인율, 위험 회피 및 위험 추구를 고려할 때, 원하는 효과와 원하지 않는 효과의 격차는 어느 정도인가?	영향을 받은 사람들의 가치를 고려할 때(예컨대, 원하는 효과와 원하지 않는 효과를 고려할 때), 원하는 효과가 원하지 않는 효과보다 클수록, 해당 중재법이 권고될 가능성은 높아진다.

기준	질문	질문 설명	세부지침
7. 필요 자원과 관련된 근거의 확실성 (근거수준)	필요 자원(비용)과 관련된 근거의 확신 정도는 어떠한가?	자원 사용 유형(예. 의약품, 입원치료)에 따라 비용 차이가 있다는 근거는 얼마나 확실한가?	
8. 필요 자원	자원(비용)은 얼마나 많이 필요한가?	자원 사용에 따른 비용 측면에서 해당 중재법과 비교 중재법은 얼마나 큰 차이가 있는가?	비용이 많이 들수록, 해당 중재법이 우선시될 가능성은 낮아진다. 반대로 많은 비용을 절약할수록, 해당 중재법이 우선시될 가능성은 높아진다.
9. 비용-효과성	해당 중재법이 비교 중재법보다 더 비용-효과적인가?	비용, 편익이 불확실성, 또는 가변성과 민감도 분석 결과와 경제성 평가의 신뢰성 및 적용 가능성을 감안할 때, 비용-효과적인 중재법인가?	효과 단위당 비용이 높을수록, 해당 중재법이 우선시될 가능성은 낮아진다.
10. 형평성	건강불평등에 미치는 영향은 어떠한가?	취약 계층(또는 해당 중재법의 절대적인 효과성 또는 문제의 중요도에 영향을 미치는 취약 계층의 기본 조건 차이)에 있어, 상대적 효과에 차이를 야기할 많나 이유가 있는가?	불평등을 완화하는 정책/프로그램이 그렇지 않은 정책/프로그램(또는 불평등을 심화시키는 정책/프로그램)에 비해 우선시될 가능성이 높다.
11. 수용성	해당 중재법은 주요 이해관계자들에게 수용될 수 있는가?	주요 이해관계자들이 편익, 위해, 비용의 분배를 수용하지 않을 가능성이 있는가? 또는 미래의 원하는 효과(편익)를 위해 단기적인 비용 또는 원하지 않는 효과를 감수할 가능성이 있는가? 주요 이해관계자들은 이들의 원하는 효과 또는 그렇지 못한 효과에 대한 가치 판단이나 윤리 문제에 대한 우려 때문에 반대할 가능성이 있는가?	주요 이해관계자른이 수용할 가능성이 낮은 방안은 권고될 가능성도 낮다. 설령 권고되더라도, 수용에 관한 문제에 대처하는 실행 전략을 수반하는 방안이 권고될 것이다. 수용 가능성은 수혜자(또는 피해자)와 지불자(또는 비용을 절감하는 주체), 그리고 편익, 역효과, 비용이 발생하는 시점(그리고 핵심 이해관계자들 가운데 방안을 무시하는 사람들의 비율-예컨대 무시하는 사람들이 많을 수 있음)에 따라 달라질 수 있다. 이해관계자들이 다음과 같이 행동하면, 수용되지 않을 수도 있다. · 편익, 위해, 비용의 분배를 수용

기준	질문	질문 설명	세부지침
			할 수 없을 때 · 미래의 원하는 효과(편익)를 위해 단기적으로 비용을 부담하거나 원하지 않는 효과를 감수하는 것을 수용할 수 없을 때 · 개인적인 이해관계 또는 타인에게 차지하는 상대적 중요도에 따라, 원하는 결과나 비용보다 원하지 않느 sruf과에 가치를 부여할 때 · 윤리적인 이유(예컨대, 자율성, 무해성, 자선 또는 정의와 같은 윤리적 원칙과 관련된 이유)
12. 수행가능성	해당 중재법은 수행될 수 있는가?	지속적으로 유지될 수 있고, 수행되는 과정에서 잠재적인 장애물도 극복할 수 있는 방안인가?	실행가능성(완수될 또는 수행될 가능성)이 낮을수록, 해당 방안이 권고될 가능성은 낮이진다(예컨대, 극복하기 어려운 장애물이 더 많을 것이다.)

2.5. 보험급여 결정을 위한 근거에 기반한 의사결정 틀

본 연구에서는 의사결정의 유형중에서도, 보험급여결정시 의사결정에 도움을 주기 신의료기술의 권고안을 마련하는 연구목적을 고려하여 보험급여 결정을 위한 EtD 틀을 추가로 살펴보았다. 보험급여 결정을 위한 EtD 틀은 검사 및 치료법, 시술 등의 새로운 보건의료 중재행위에 대한 보험급여여부의 의사결정을 지원에 도움을 주기 위한 틀로써, 7가지 구성요소(질병부담, 이득 및 위해, 가치 및 선호, 자원사용, 형평성, 정책제도내의 수용가능성, 보건의료체계내의 실행가능성)를 체계적으로 평가하는 도구(그림)이다. 동 EtD framework는 기존의 틀을 유지하면서, 사용중에 있는 대체기술의 제한점과 정책적 및 보건시스템 범위내에서의 수용성과 실현가능성을 고려할 수 있도록 수정하였다.

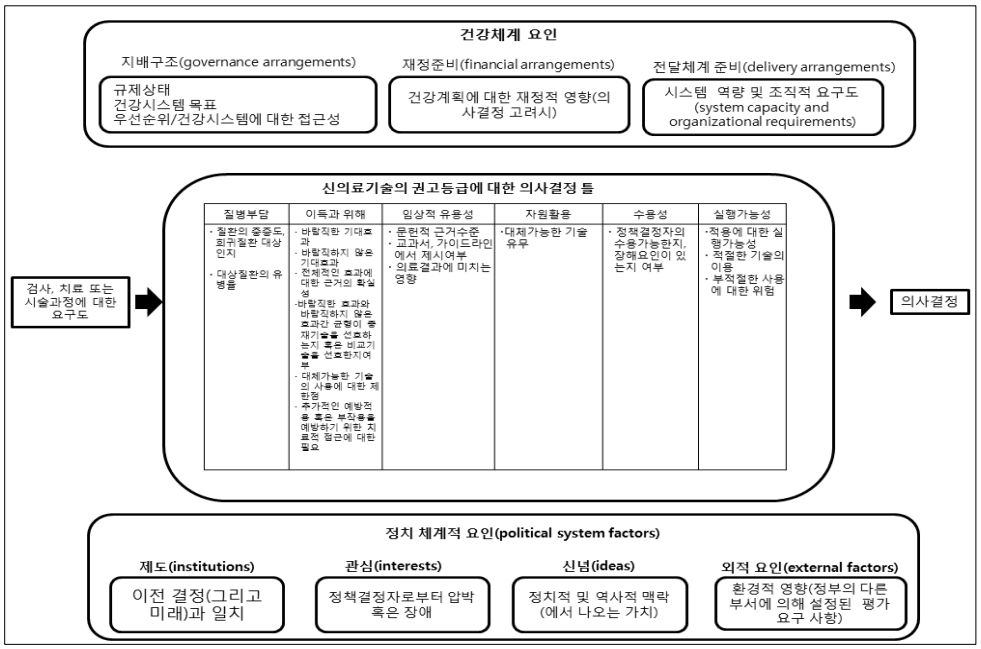

그림 8. 의료중재에 대한 통합된 의사결정 framework

(출처 : Morgan et al., 2018)[46]

46) Morgan RL, Kelley L, Guyatt GH, Johnson A, Lavis JN. Decision-making frameworks and considerations for informing coverage decisions for healthcare interventions: a critical interpretive synthesis.J Clin Epidemiol. 2018 Feb:94:143-150. doi: 10.1016/j.jclinepi.2017.09.023. Epub 2017 Oct 6.

본 연구에서는 Parmelli(2017)에서 보고한 EtD 틀을 활용한 예제를 인용하여 제시하였다[47].

가, 질문의 작성

보험급여 의사결정을 수행하고자 하는 주제에 대한 질문을 PICO 형식, 즉 사용대상 / 중재/ 비교/ 주요 의료결과/ Setting / Perspective별로 상세하게 설정해야 한다.

표 37. Coverage를 위한 EtD framework (질문섹션)

질문
Z condition을 위해, X의 사용을 보험해 줄 것인가?
사용대상 (population) : 배경(Background) :
중재기술 (Intervention) :
비교기술 (Comparison) :
주요 의료결과 (Main outcome) :
Setting :
Perspective :

출처 : Dahm et al (2017) ; Elena Parmelli et al (2017)

<예제 : 시나리오>	
무증상 남성에서 전립선암의 선별검사에 대한 보험적용을 중단해야 하는가?	
Patients : 50세 이상의 무증상 남성 Intervention : PSA 선별검사 제공 Comparison : 선별검사 시행하지 않음 Main Outcome : 모든 원인 사망률, 전립선암 사망률, 삶의 질, 위해 Setting : 이탈리아의 국민건강보험 Perspective : 지역보건당국	Background : 전립선암은 남성에게 이환율/사망률의 주요원인임. 조기에 증상징후를 유발하는 경우는 드묾. 무증상 남성의 효과적인 조기 진단 및 치료 전략은 잠재적으로 많은 남성에게 큰 이익을 줄 수 있음. 선별검사는 조기에 암을 진단하여 성공적인 치료를(생존, 삶의 질 향상) 증가시키는 것을 목표로 함. 그러나 많은 남성들이 다른 원인의 암으로 사망할 때까지 무증상 전립선암으로 생존함. 증상이나 사망을 일으키지 않는 암을 진단하는 것은 과잉진단이라고 함. 과잉진단의 결과에는 불필요한 확인으로 인한 부정적 영향, 불필요한 검사 및 치료의 고통, 기회비용의 낭비 등이 포함됨

47) Parmelli E, Amato L, Oxman AD, Alonso-Coello P et al., GRADE Evidence to Decision(ETD) Framework for coverage Decisions. International Journal of Technology Assessment in Health Care 2017; 33: 176-182.

나, 평가

보험급여 결정을 위한 EtD 틀은 GRADE EtD 구조적 틀을 기반으로 하고 있으며, 질병부담(burden of disease), 이득과 위해, 가치 및 선호, 자원 사용, 형평성, 수용 가능성 및 실현 가능성이 포함되었다. EtD 틀을 활용한 평가는 보험결정에 영향을 미칠수 있는 12가지 요소(기준)(표38)을 이용하여, 각 항목별 이용가능한 최선의 근거에 의한 판단과, 기타 고려사항을 통해서 의사결정하여야 한다.

표 38. 근거에 기반한 의사결정 틀(Evidence to Decision Frameworks)

기준(질문)		판단	연구근거	추가고려사항
1. 문제	관련 문제는 우선사항인가?	○ 잘 모르겠음 ○ 다양함 ..... ○ 우선사항이 아님 ○ 아마도 우선사항이 아님 ○ 아마도 우선사항임 ○ 우선사항임	예 위험비:사망률, 이환율, 발생률 등	
2. 가치	주요 결과에 대한 가치판단에 중요한 불확실성이나 가변성이 존재하는가?	○ 관련되어 알려진 바 없음 ..... ○ 중요 불확실성 또는 가변성이 있음 ○ 중요 불확실성 또는 가변성이 있을 수 있음 ○ 아마도 중요 불확실성 또는 가변성이 없을 것임 ○ 중요 불확실성 또는 가변성이 없음	주요 관심결과의 상대적 중요도 및 가치	
3. 효과의 확신도 (근거수준)	효과와 관련된 근거의 확실성은 전반적으로 어떠한가?	○ 관련된 연구 없음 ..... ○ 매우 낮음 ○ 낮음 ○ 보통 ○ 높음	결과요약표: 연구결과의 변화가 의사결정에 영향을 미칠 가능성	
4. 원하는(긍정적) 효과	원하는 효과의 크기는 어느 정도인가?	○ 관련된 연구 없음 ○ 다양함	결과요약표	

기준(질문)		판단	연구근거	추가고려사항
		····· ○ 경미함 ○ 적음 ○ 보통 ○ 많음		
5. 원하지 않는 (부정적) 효과	원하지 않는 효과의 크기는 어느 정도인가?	○ 관련된 연구 없음 ○ 다양함 ····· ○ 많음 ○ 보통 ○ 적음 ○ 경미함	결과요약표	
6. 효과의 비교	원하는 효과와 원하지 않는 효과를 모두 고려했을 때, 해당 중재법이 비교 중재법보다 더 선호되는가?	○ 관련된 연구 없음 ○ 다양함 ····· ○ 비교 중재법이 더 선호됨 ○ 아마도 비교 중재법이 더 선호됨 ○ 아마도 해당 중재법이 더 선호됨 ○ 해당 중재법이 더 선호됨	[가치] 주요 관심결과의 상대적 중요도 및 가치 [효과] 결과요약표	
7. 필요자원과 관련된 근거의 확실성(근거수준)	필요 자원(비용)과 관련된 근거의 확신 정도는 어떠한가?	○ 관련된 연구 없음 ····· ○ 매우 낮음 ○ 낮음 ○ 보통 ○ 높음		
8. 필요 자원	자원(비용)은 얼마나 많이 필요한가?	○ 잘 모르겠음 ○ 다양함 ····· ○ 많은 비용이 필요함 ○ 어느 정도 비용이 필요함 ○ 비용 변화의 차이 거의 없음 ○ 어느 정도 비용이 절감됨 ○ 많은 비용이 절감됨	예) 보험수가;	

기준(질문)		판단	연구근거	추가고려사항
9. 비용-효과성	해당 중재법이 비교 중재법보다 더 비용-효과적인가?	○ 관련된 연구 없음 ○ 다양함 …… ○ 비교 중재법이 더 비용-효과적임 ○ 아마도 비교 중재법이 더 비용-효과적임 ○ 비용-효과성의 차이 거의 없음 ○ 아마도 해당 중재법이 더 비용-효과적임 ○ 해당 중재법이 더 비용-효과적임	예) 질보정수명(QALY);	
10. 형평성	건강불평등에 미치는 영향은 어떠한가?	○ 잘 모르겠음 ○ 다양함 …… ○ 건강불평등이 증가함 ○ 아마도 건강불평등이 증가함 ○ 아마도 건강불평등이 감소함 ○ 건강불평등이 감소함		
11. 수용성	해당 중재법은 주요 이해관계자들에게 수용될 수 있는가?	○ 잘 모르겠음 ○ 다양함 …… ○ 수용될 수 없음 ○ 아마도 수용될 수 없음 ○ 아마도 수용될 수 있음 ○ 수용될 수 있음		
12. 수행가능성	해당 중재법은 수행될 수 있는가?	○ 잘 모르겠음 ○ 다양함 …… ○ 수행될 수 없음 ○ 아마도 수행될 수 없음 ○ 아마도 수행될 수 있음 ○ 수행될 수 있음		

재인용: 한국보건의료연구원, 2015[48], 원저: iEtD 홈페이지[49]

48) 김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015. 324-6p.

1) 문제의 우선순위(Priority of the problem)

치료 대상이 되는 문제의 우선순위에 따라 보험결정이 이루어진다. 우선순위는 문제의 심각성, 시급성, 국가의 보건계획 상 우선순위에 있는 사항인지, 문제에 많은 사람들이 영향을 받는지 등 질병부담이 고려사항이 될 수 있다. 이를테면, 경미한 질병보다는 중증도가 높은 질병이 보험급여 대상에서 우선순위가 높을 것이며, 그 외 정책적인 요구도가 있는 경우, 문제로부터 영향을 받은 대상자가 많을수록 급여 결정에 있어 우선순위로 삼을 수 있다.

Q. Is the problem a priority?
Judgment

☐	☐	☐	☐	☑	☐
Don't know	Varies	No	Probably No	Probably Yes	Yes

Research evidence
– 전립선암은 고소득국가의 남성 사망원인 중 세번째로, 가장 흔히 진단되는 암임. 고연령이 위험인자로, 75%이상의 전립선 암이 65세 이상의 남성에서 진단됨
– 대부분의 전립선 암 환자는 증상이 없고, 일반적인 검사로 종양이 발견됨. 양성전립선폐쇄로 인한 하부요로증상은 노인 남성에서 흔히 발생하며, 전립선특이항원(PSA)의 농도를 증가시킬수 잇으나, 전립선암 발생 증가와는 관련이 없음. 대부분의 남성에서 전립선암은 느리게 진행하며, 평생동안 임상징후를 나타내지 않음

2) 기대하는 효과(Desirable effects) / 기대하지 않는 효과(Undesiable effects)

중재로부터 기대하는 효과(건강 및 기타 이득) 또는 기대하지 않는 효과(건강 및 기타 위해)의 크기를 고려하여 평가한다. 즉, 예상되는 기대하는 편익 결과와 기대하지 않는 부작용의 중요성/심각성 및 이로부터 영향을 받는 환자의 수가 고려대상이다. 기대하는 효과의 크기가 크면 클수록, 기대하지 않는 효과 크기가 작을 수록 권고결정이 내려진다.

Q. How substantial are the desirable anticipated effects?
Judgment

☐	☐	☐	☑	☐	☐
Don't know	Varies	Trivial	Small	Moderate	Large

49) iEtD 홈페이지

Q. How substantial are the undesirable anticipated effects?

Judgment

☐	☐	☑	☐	☐	☐
Don't know	Varies	Large	Moderate	Small	Trivial

Research evidence

– 근거요약표(SoF) 확인

Outcome	Plain language summary	Absolute Effect Without PSA	 With psa	Relative effect (95% CI) N° of participants & studies	Certainty of the evidence (GRADE)
All-cause mortality	*No statistically significant difference between men randomised to screening and control groups.*	210 per 1000	210 per 1000 Difference 0 more per 1000 Men (95% CI: 1 less to 1 more per 1000 Men)	RR 1 (0.96 to 1.03) Based on data from 294856 Men in 4 studies	⊕⊕⊕○ Moderate
Prostate cancer mortality	*No statistically significant difference between men randomised to screening and control groups.*	7 per 1000	7 per 1000 Difference 0 more per 1000 Men (95% CI: 1 less to 1 more per 1000 Men)	RR 1 (0.86 to 1.17) Based on data from 341342 Men in 5 studies	⊕⊕⊕○ Moderate
Prostate cancer diagnosis	*Diagnosis of prostate cancer is significantly greater in men randomised to screening compared to those randomised to control.*	68 per 1000	88 per 1000 Difference 20 more per 1000 Men (95% CI: 1 to 42 more per 1000 Men)	RR 1.3 (1.02 to 1.65) Based on data from 294856 Men in 4 studies	⊕⊕○○ Low

3) 효과에 대한 근거 확실성(Ceratainty of the Evidence of Effects)

근거의 확실성 평가는 체계적 문헌고찰 과정에서는 문헌에서 도출된 효과추정치가 정확하다는 것을 얼마나 신뢰할만한가를 반영하는 것이라 한다면, 의사결정 과정에서는 효과추정치가 특정 의사결정이나 권고를 뒷받침하기에 충분하다는 것에 대한 확신의 정도를 의미한다.

중재의 주요 의료결과별로 상대적인 중요성(가치)을 판단 후, 문헌에서 도출된 결과 요약표를 통해 중재 효과에 대한 근거의 확실성의 정도를 평가한다.

효과에 대한 근거 확실성이 낮은 경우 보험급여에 영향을 미친다. 이를테면, 기대한 효과가 불확실한 경우 중재를 급여하지 않는 것으로 결정한다.

중재가 유망한 기술이지만 중요한 불확실성이 있다면, 불확실성의 정도와 다른 요인들을 고려하여 잠재적 부작용 효과를 모니터링하면서 급여를 할지, 연구의 근거생성을 조건으로 한 조건부 급여를 할지, 근거가 축적될 때까지 급여하지 않을지를 결정하여야 한다.

Q. What is the overall certainty of the evidence of effects?
Judgment

☐	☐	☐	☑	☐
No included studies	Very low	Low	Moderate	High

Research evidence
- 상기 평가요약표(SoF) 참고
- 전립선암 선별검사에 대한 코크란 리뷰 근거에 의해, 선별검사는 전체 원인 사망률 및 암 특이 사망률에 유의한 효과 차이는 없는 것으로 나타남(중증도 근거수준)

4) 가치에 대한 불확실성(Uncertainty about how much the people affected value the main outcome)

주요 의료결과는 임상적 권고와 마찬가지로 중재기술의 보험급여 여부를 결정하는데 중요한 요인이다. 의료결과의 상대적인 중요성에 따라, 환자들이 중재기술에 얼마나 많은 영향을 받는지, 환자의 가치에 대해 중요한 불확실성이 있는지 평가해야 한다. 즉, 환자 입장에서 의료결과가 주는 가치가 있다고 판단할만한 근거가 있는지 또는 다른 의사결정을 내릴 수 있을 정도의 가치를 변동시킬 근거가 있는지 고려해야 한다. 예를 들면, 심각한 부작용이 있는 항암화학요법의 경우 항암화학요법의 기대하는 효과(생존 연장)와 기대하지 않는 효과(메스꺼움, 탈모, 감염의 위험증가와 같은 부작용)간의 상대적인 중요성에 대한 가치 판단이 요구된다.

환자의 가치에 중요한 불확실성이 있는 경우 중재의 보험급여 결정을 연기하거나 연구 목적에서의 조건부 급여로 결정할 수 있다. 임상적 권고와는 다르게, 얼마나 많은 환자들이 주요 결과가 가치있다고 생각하는지 그 다양성(variability)은 보험급여 결정과 관련이 없다. 즉, 기대하는 효과가 기대하지 않은 효과보다 더 가치있다고 생각하는 환자들에서는 중재를 보험 급여하는 것이 정당화 될 수 있으나, 그렇게 생각하지 않는 사람들은 중재의 사용을 선택하지 않을 수도 있기 때문이다.

Q. Is there important uncertainty about how much people value the main outcomes?
Judgment

☐	☐	☑	☐
Important uncertainty	Possibly important uncertainty	Probably no important uncertainty	No important uncertainty

Research evidence
- 전립선암 검진에 대한 남성 선호도를 확인하기 위한 연구(de Bekkeer-Grob 2012)에 따르면, 남성은 생검이나 불필요한 치료 부담을 경감시키기 보다는 전립선암 관련 사망 위험을 줄이고자 하는 것으로 나타남
- 전립선암의 1차 치료후 건강 관련 삶의 질의 결정요인 및 치료결과의 만족도에 미치는 결정요인 효과를 평가한 연구(Sanda 2008)에서, 각 전립선암 치료는 비뇨기, 성기능, 내장 및 호르몬 기능과 관련한 삶의 질 영역 변화에 영향이 있는 것으로 나타남. 이러한 변화는 환자 및 배우자의 치료결과의 만족도에 영향을 미침

5) 효과의 균형(Balance of effects)

기대한 효과가 기대하지 않은 효과 결과 사이에서, 상대적 중요성의 크기를 평가한다. 기대하지 않는 결과보다 기대한 효과가 클수록 영향을 받는 환자들의 가치를 고려하면, 중재에 대해 긍정적인 권고를 내릴 가능성이 높다.

Q. Does the balance between desirable and undesirable effects favor the intervention or the comparison?
Judgment

☐	☐	☑	☐	☐	☐	☐
Don't know	Varies	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention

Research evidence
- 상기 평가요약표(SoF) 참고

6) 자원의 사용 (Resource Use)

경제성 평가는 급여결정에 중요한 요인이다. 중재비용이 많이 들고, 비용효과가 낮을수록 급여될 가능성은 낮다. 표준치료와 비교하여, 중재의 점진적 비용효과성에 대해 중요한 불확실성이 있는 경우, 보험 의사결정은 근거가 더 축적될 때까지 보류하거나, 효과와 지출을 모니터링하면서 중재기술을 급여하도록 결정 할 수 있다.

Q. (자원사용 크기) How large are the resource requirements (costs)?
Judgment

☐	☐	☐	☑	☐	☐	☐
Don't know	Varies	Large costs	Moderate costs	Negligible costs or savings	Moderate savings	Large savings

Research evidence

Age	Total population	N° of patients	N° of PSA performed	% patients	Single cost €	Total costs €
50-59	36 781	6 302	8 754	17.1	7.41	64 867
60-69	26 975	9 058	14 631	33.6	7.41	108 416
70-79	22 461	11 133	20 275	49.6	7.41	150 238
>79	13 038	5 929	10 716	45.5	7.41	79 406
≥ 50	**99 255**	**32 422**	**54 376**	**32.7**	7.41	**434 781**

- PSA 선별검사 관련 비용 : 생검(50€ each), specialists's visits (30€ each), 치료비(3000€ each), 합병증 치료(200€ each)
- 자료는 2013년 로마E 이탈리아 지역 보건당국(주민 537,002명)

Q. (자원사용에 대한 근거수준) What is the certainty of the evidence of resource requirements (costs)?

Judgment

☐	☐	☐	☑	☐
No inclued studies	Very low	Low	Moderate	High

Research evidence

- 비용에 관한 데이터는 실제 환자 정보를 분석한 지역 보건당국 DB에서 추출됨

Q. (비용-효과성) Does the cost-effectiveness of the intervention favor the intervention or the comparison?

Judgment

☐	☐	☑	☐	☐	☐	☐
Don't know	Varies	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention

Research evidence

- (Shteynshyuger 2011) PSA 평가 및 치료하는데 드는 PSA 선별검사의 평생 비용은 5,227,306 달러였음. 선별검사가 전립선 암 특이 사망률을 감소와 같이 전체 사망률 감소율을 달성한다면, 1 인당 생존연수 당(per life-year saved) 262,758 달러의 비용이 듦. 연구저자는 $100,000/LYS의

비용효과성 임계값을 사용하여 PSA 선별검사가 비용가치가 높지 않음을 제안함
- (Shin S 2014) 한국 50-74세 남성의 PSA 검진프로그램의 채택에 대한 비용효용 분석 결과, PSA 검진은 비용 효과적이지 않음. PSA 검사의 순이득은 매우 낮게 평가됨. ICER는 QALY 당 약 9천4백만원(약 76,140달러)이었음
- (Pataky R 2014) British Columbia 인구집단에서의 PSA 검사의 비용 효과성 분석시, 효용(utility) 보정 후에 모든 선별검사 전략은 QALY가 손실되었음. utility estimate는 매우 민감하였음

7) 형평성 (Equity)

중재에 대한 형평성 있는 접근은 보험 급여 결정을 위한 중요한 고려사항이다. 불평등을 감소하는 정책 또는 프로그램은 불평등한 정책보다는 우선순위가 높다.

Q. What would be the impact on health equity?
Judgment

☐	☐	☐	☐	☑	☐	☐
Don't know	Varies	Reduced	Probably reduced	Probably no impact	Probably increased	increaed

Research evidence
- No evidence found

8) 수용성 (Acceptability)

보험 급여 의사결정시 수용성에 대한 고려사항은 임상적 권고와 유사하며, 윤리적 가치 뿐 아니라 이득, 위해 및 비용의 배분에 대한 고려사항까지 포함한다. 이는 보험결정에 영향을 받는 이해관계자들(이를테면, 직/간접적으로 보험급여를 통해 혜택을 받는 사람, 보험혜택을 받지는 않지만 보험료 및 세금을 통해 중재비용을 분담하는 사람, 보건의료 전문가, 의료기술 제조자, 중재기술의 이익을 대표하는 조직)로부터의 중재에 대한 수용성을 의미한다. 중재에 대한 수용성이 낮으면, 권고를 덜 하게 되거나 권고하게 될 경우에라도 수용성에 대한 우려를 해결하기 위한 실행전략을 포함시킬 가능성이 높다.

일부 이해관계자들에 의해, 다음과 같은 사유로 수용되지 않을 수 있음
- 이득, 위해 및 비용의 분배가 수용되지 않음
- 기대하는 효과(이득)를 위해, 단기간에 발생되는 비용 및 기대하지 효과가 수용되지 않음

- 중재로부터 기대하는 결과 및 비용보다 기대하지 않은 결과에 가치(상대적인 중요성)를 더 많이 부여하는 경우
- 도덕적으로 승인되지 않는 경우

Q. Is the intervention acceptable to key stakeholders?
Judgment

☐	☐	☐	☐	☑	☐
Don't know	Varies	No	Probably No	Probably Yes	Yes

Research evidence
– No evidence found
Additional consideration
– 50세 이상의 남성을 대상으로 한 PSA 선별검사는 이탈리아에서 널리 사용됨. 보험급여를 중지하는 것은 일부에게 수용되지 않을수 있음
  • 이미 선별검사를 받은 남성
  • 과거 routine 검사였기 때문에 선별검사를 요청한 남성
  • 전립선암의 가족력이 있는 남성

9) 실행가능성 (Feasibility)

보험급여결정을 위한 실행가능성 고려사항은 임상적 권고를 위한 고려사항과는 다르다. 보험급여를 위한 고려사항은 중재에 대한 급여, 제한적 급여 또는 비급여를 시행하는데 실행가능성에 중점을 둔다.

중재에 대한 수요증가를 충족시키는 수용력(capacity) 또한 중요한 고려사항이다. 예를 들면, 대장내시경 검사를 제공할 수 있는 보건의료시스템 수용력은 대장암 진단을 위해 대장내시경 선별검사를 급여할지 여부를 결정 할 때 중요한 고려사항이 될 수 있다. 수용가능성(acceptability)과 유사하게, 실행가능성 고려사항은 중재급여 여부 결정을 변경하지 않을 수 있지만, 결정이 실제 보급되고 구현되는 상황을 고려하여 심사숙고되어야 한다.

Q. Is the intervention feasible to implement?
Judgment

☐	☐	☐	☐	☑	☐
Don't know	Varies	No	Probably No	Probably Yes	Yes

Research evidence
- No evidence found
Additional consideration
- 임상의는 잠재적으로 무증상 남성에 대해 PSA 검사를 계속할 것이며, 검사의 부정확한 사유 제공 또는 환자의 비용부담(pay out-of-pocket) 제안을 할 것임

평가항목들에 대한 각 평가를 수행한 후, 이를 종합하여 하나의 평가요약본으로 나타낸다.

	Judgement							Implications
Problem								
Desirable Effects								
Undesirable Effects								
Certainty of evidence								
Values								
Balance of effects								
Resources required								
Certainty of evidence of required resources								
Cost effectiveness								
Equity								
Acceptability								
Feasibility								

다, 결론 도출(Drawing Conclusion)

보험결정을 위한 EtD framework은 5가지 보험결정의 유형, 즉 급여 미적용(no coverage), 근거 생산 조건부 급여(coverage with evidence development(in the context of research)), 가격협상을 통한 급여(coverage with price negotiation), 제한적 급여(restricted coverage), 급여(full coverage)으로 구분되어 결정하도록 되어 있다. 중재를 평가하는 의사결정기준에서 하나 이상을 충족하는 경우 보험결정을 내릴 수 있으며, 이는 의사 결정마다 달라질 수 있다.

임상근거 생성을 조건으로 하는 연구맥락에서의 조건부 급여는 중재의 효과에 대해 중요한 불확실성이 존재할 때 결정될 수 있다. 새롭고 유망한 중재기술이지만, 임상에 바로 적용하기가 어려운 경우가 이에 해당된다.

가격협상을 통한 급여는 자원 사용 또는 비용효과성에 대한 기준을 충족하지 못하는 새롭고, 효과적인 약물에서 주로 결정이 이루어진다. 가격협상은 제조업체와 지불자 간의 위험분담계약(risk-sharing agreements)을 포함하고 있다.

제한적 급여는 효과 또는 비용효과가 있는 특정 하위 집단에 대해 중재를 사용하도록 할 때 결정할 수 있다. 제한적 급여는 특정 환자에게 의료기술이 권고되는 조건부 권고와 유사하다.

중재의 급여를 결정하는 것은 권고와 동일하지 않다. 즉, 중재를 급여한다는 것은 강력한 임상적 권고 또는 약한 임상적 권고가 모두 해당될 수 있다.

보험 결정을 위해서는, 보험 적용이 제공되는 중재의 사용, 부적절한 사용 및 비용을 모니터링하는 것이 중요하다. 부적절한 사용이 우려되는 경우에는 레지스트리 또는 다른 수집된 데이터를 사용하여 모티터링 할 수 있으나, 이를 잘 실행하기 위해서는 신뢰할 수 있는 데이터가 수집되어야 하고, 관찰된 수치들이 사전에 정의된 최적의 기준과 비교 가능해야 하며, 가능하다면 보험결정에 영향을 받는 이해관계자의 동의를 받아야 한다.

Conclusions					
Should x vs. y be used for z? Type of decision	Do not cover	Cover with evidence development	Cover with price negotiation	Cover with restrictions	Cover
	○	○	○	○	○
Decision					
Justification					
Restrictions					
Implementation considerations					
Monitoring and evaluation					
Research priorities					

Type of decision

☐	☐	☐	☐	☐
do not cover	cover with evidence development	cover with price negotiation	restricted coverage	cover

Research evidence

- None

Justification

- 50세 이상의 무증상 남성에서의 PSA 선별검사 제공은 사망률, 삶의 질 측면에서 이득이 없으며, 출혈, 멍, 단기 불안, 과잉진단 및 과잉치료 등을 포함하는 기대하지 않는 효과에 미치는 영향이 있음

Restriction

– No restrictions

Implementation consideration

– 환자 정보가 제공되어야 하고, 선별검사를 하지 않는 이유에 대해 해당 남성에게 명확하게 전달되어야 함

Monitoring and evaluation

– 무증상 남성에서의 PSA 선별검사의 사용은 모니터되어야 함

3. 신의료기술평가자료의 현황 분석 결과

3.1. 신의료기술 평가 전체 신청현황

2007년 신의료기술평가제도 도입 이후부터 2017년 12월 31일까지 신의료기술평가의 총 신청 건수는 2,295건이었다. 이 중 기존기술로 심의된 870건의 기술을 제외하면 신의료기술평가가 수행된 건은 총 1,425건이었으며, 이를 평가기술 및 적응질환별로 구분하면 총 1,643건으로 확인되었다.

평가 신청 기술을 기술유형별로 구분하여보면, 가장 많이 신청된 기술유형은 체외진단검사로 전체 신청건수의 54%를 차지하였으며, 그 다음으로 처치 및 시술이 33%, 기타검사가 10%, 영상검사가 3% 순이었다. 평가 신청 기술을 평가결과별로 구분한 결과에서는 평가가 수행된 1,643건 중 신의료기술로 인정된 건이 49%(800건)이었으며, 이 외 연구단계기술이 23%(373건), 조기기술이 29%(470건)으로 확인되었다.

신의료기술로 인정된 기술을 기술유형별로 구분한 결과에서는 체외진단검사가 64%(508건)으로 가장 많은 비율을 차지하였으며, 그 다음으로 처치 및 시술이 12%(192건), 기타검사가 4%(65건), 영상검사가 2%(35건) 순이었다. 반면, 연구단계기술로 평가된 기술 중 가장 많은 비중을 차지하는 기술유형은 처치 및 시술(11%, 176건)로 나타났으며, 조기기술로 평가된 기술에 있어서는 신의료기술 인정 기술들과 동일하게 체외진단검사(14%, 236건)의 비율이 가장 높았다.

기술유형별로 신의료기술평가 결과를 구분한 결과에서는 처치 및 시술의 경우에는 신의료기술로 인정된 비율과 연구단계기술, 또는 조기기술로 심의된 비율이 32 ~ 35%로 유사한 수준을 보였다. 반면, 체외진단검사와 영상검사, 기타검사의 경우에는 신의료기술로 인정된 기술의 비율이 각각 58%, 63%, 40%로 연구단계기술, 또는 조기기술로 심의된 결과보다 더 높았다.

신의료기술평가 신청 건에 대한 평가결과 및 기술유형별 전체 현황의 자세한 내용은 다음의 표 39, 40 및 그림 10, 그림 11과 같다.

표 39. 평가결과 및 기술유형별 신의료기술평가 전체 신청 현황

단위: 건(%)

구분	처치 및 시술	체외진단검사	영상검사	기타검사	합계
신의료기술	192(24)	508(64)	35(4)	65(8)	800(49)
연구단계기술	176(47)	138(37)	19(5)	40(11)	373(23)
조기기술	176(37)	236(50)	2(0)	56(12)	470(29)
전체	544(33)	882(54)	56(3)	161(10)	1643(100)

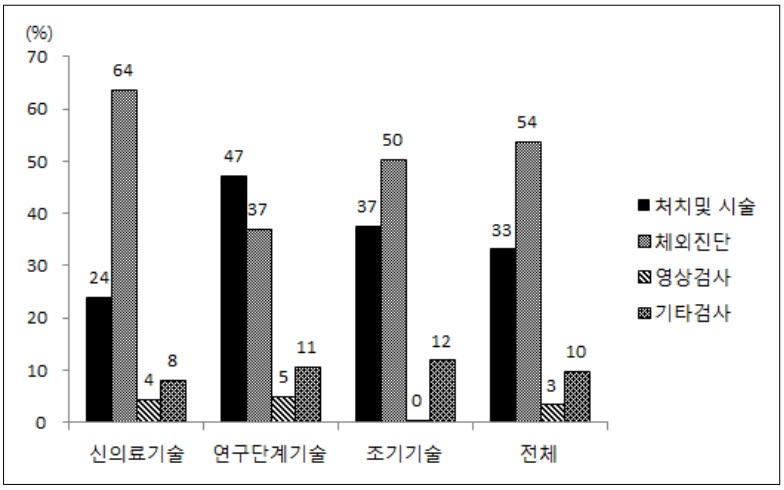

그림 9. 평가결과별 기술유형 분포

표 40. 기술유형별 신의료기술 평가 결과 분포

구분	처치 및 시술	체외진단	영상검사	기타검사	합계
신의료기술	35%	58%	63%	40%	49%
연구단계기술	32%	16%	34%	25%	23%
조기기술	32%	27%	4%	35%	29%
합계	100%	100%	100%	100%	100%

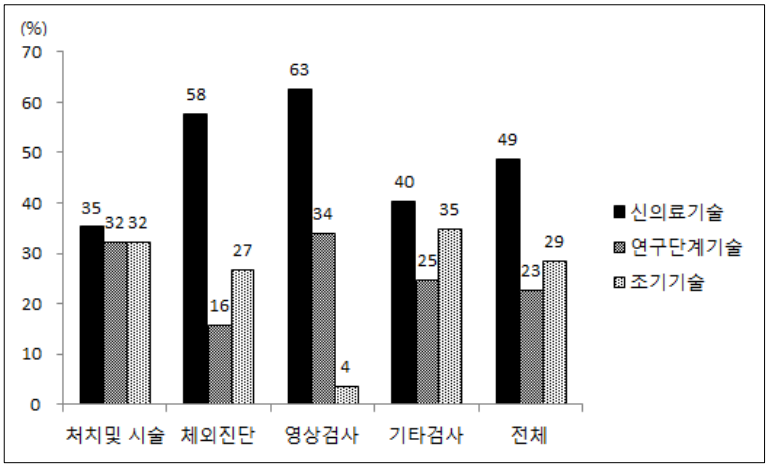

그림 10. 기술유형별 신의료기술 평가 결과 분포

3.2. 신의료기술평가 신청 후 평가가 수행된 기술의 관련 현황

신의료기술평가가 신청 된 후 체계적 문헌고찰이나 전문 소위원회의 논의를 거쳐 신속으로 평가된 기술 총 1,279건에 대해 현황 분석을 수행하였다. 동 평가에서는 평가대상여부 심의단계에서 전문소위원회 검토 없이 신의료기술평가위원회에서 조기기술로 심의된 기술들은 신의료기술 평가대상으로서의 심층평가가 수행되지 않은 기술들이므로 분석에서 제외하였다.

3.2.1. 전체 분석결과

가, 평가방법에 따른 신의료기술평가 결과

신의료기술평가의 평가 방법에 따라 심의결과의 현황을 분석하였다. 신의료기술평가의 방법은 크게 체계적 문헌고찰 방법론을 활용한 평가와 신속검토를 통한 신속평가 두 가지 방법으로 구분할 수 있다. 대부분의 경우 평가를 위해 근거에 대한 충분한 검토가 필요하므로 체계적 문헌고찰 방법론을 활용하여 평가를 수행하나, 다양한 이유에 의해 신속검토가 가능한 안건이거나 신속하게 평가가 이루어져야하는 경우에는 체계적 문헌고찰을 통한 방법 대신 신속평가를 수행한다. 이번 분석에서 신속평가가 수행된 기술은 다음의 세 가지 기준에 부합하는 경우로 정의하였다.

i) 평가대상여부 심의단계에서 유전자 전문 검사 소위원회 및 검사전문 소위원회 논의를 거쳐 평가한 기술
ii) 교과서 및 가이드라인 등의 관련 내용 검토를 통해 담당 소위원회에서 신속하게 평가가 이루어진 기술
iii) 소위원회의 평가를 거치지 않고 바로 신의료기술평가위원회에서 신의료기술로 인정되었던 기술

분석결과, 전체 1,239건의 평가기술 중 체계적 문헌고찰로 평가가 이루어진 기술은730건(57%), 신속검토고찰을 통해 평가가 이루어진 기술은 549건(43%)으로 전체 평가에 있어서는 두 방법이 비슷한 수준으로 수행된 것으로 확인되었다. 하지만 이를 기술유형별로 구분하였을 때는 기술유형에 따라 수행된 평가방법에 차이가 있어, 처치 및 시술인 경우 95%, 영상검사의 경우 96%, 기타검사의 경우 91%의 기술이 체계적 문헌고찰 방법으로 평가되었지만 체외진단검사의 경우에는 33%의 기술만 체계적 문헌고찰 방법으로 평가되고 67%의 기술은 신속평가가 수행되어 다른 유형의 기술들과 평가방법에 차이가 있었다. 기술유형별 평가방법에 따른 신의료기술평가 결과에 대한 자세한 내용은 표 41, 그림 12과 같다.

표 41. 기술유형별 평가방법에 따른 신의료기술 평가 결과

단위: 건(%)

구분	처치 및 시술	체외진단검사	영상검사	기타검사	전체
체계적 문헌고찰	335(95)	254(33)	52(96)	89(91)	730(57)
신속평가	17(5)	521(67)	2(4)	9(9)	549(43)
전체	352(100)	775(100)	54(100)	98(100)	1,279(100)

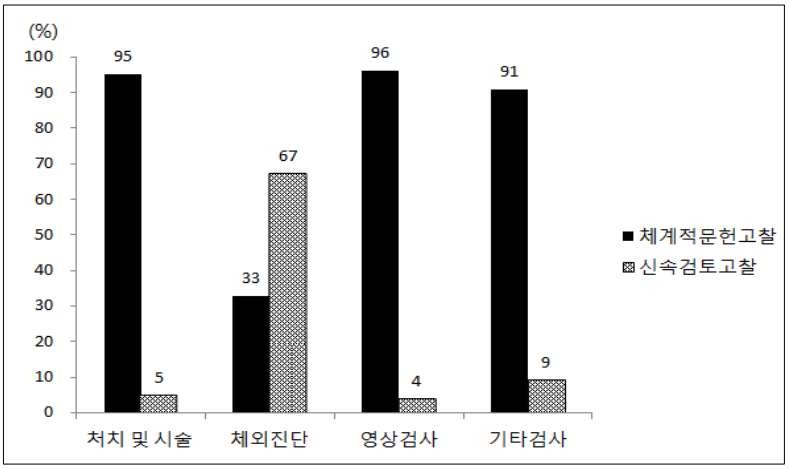

그림 11. 기술유형별 평가방법에 따른 신의료기술 평가결과 분포

나, 평가결과별 근거수준 유형

신의료기술평가가 수행된 기술 1,279건의 평가결과별 근거의 수준의 분포를 분석하였다. 신의료기술로 인정된 기술과 연구단계기술을 구분하지 않은 전체 결과에서는 근거의 수준을 미수행한 비율이 42%로 가장 높게 나타났다. 근거의 수준을 보고한 기술들에서는 근거의 수준 D로 평가된 기술의 비율이 26%로 가장 높았으며, 그 다음으로는 근거의 수준 C로 평가된 기술의 비율이 23%로 보고되었다. 근거의 수준 B와 A의 비율은 각각 8%와 2%로 보고되어 근거의 수준 C와 D에 비해 보고된 비율이 많이 낮았다.

신의료기술로 인정된 경우에서는 전체 분석 결과와 동일하게 근거의 수준 평가를 미수행한 경우가 43%로 가장 높았으나, 근거의 수준을 평가한 경우에서는 근거의 수준 C의 비율이 26%로 가장 높고 그 다음으로 근거의 수준 D 20%, B 10%, A 2%로 나타나 전체 분석 결과와 약간의 차이가 있었다.

연구단계기술의 경우에는 전체 분석 결과와 동일하게 근거의 수준 평가를 수행한 경우가 40%로 가장 높았고, 근거의 수준을 평가한 경우에서는 D 35%, C 18%, B 5%, A 2%로 전체 분석 결과와 동일한 순서를 나타내었다.

표 42. 평가결과별 근거의 수준 유형

단위: 건(%)

구분	근거의 수준				미수행	합계
	A	B	C	D		
신의료기술	14(2)	79(10)	208(26)	158(20)	341(43)	800(100)
연구단계기술	8(2)	25(5)	86(18)	169(35)	191(40)	479(100)
전체	22(2)	104(8)	294(23)	327(26)	532(42)	1,279(100)

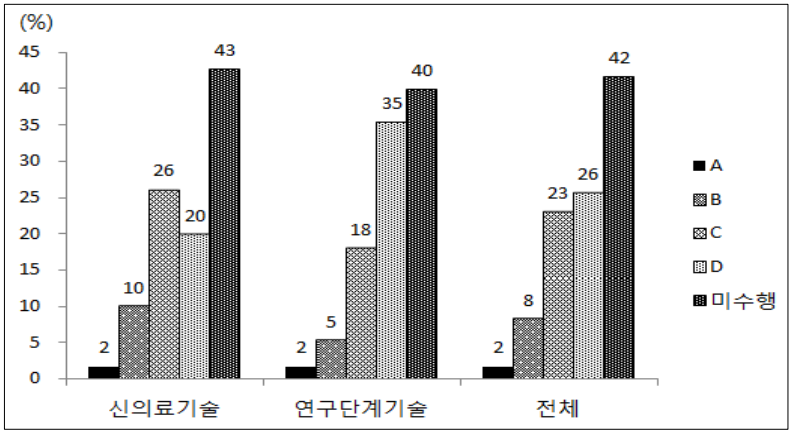

그림 12. 평가결과별 근거의 수준 유형

기술 유형별 신의료기술평가 결과에 따른 근거의 수준 평가 결과는 다음과 같다. ⅰ) 처치 및 시술의 경우 전체적으로 근거의 수준 D가 48%로 가장 높았으며 평가결과별로 구분해서 보았을때도 신의료기술, 연구단계기술 각각 근거의 수준 D가 39%, 58%로 가장 높은 비율을 차지하였다. ⅱ) 체외진단검사에서는 전체적으로 근거의 수준을 미수행한 경우가 63%로 가장 높았으며, 평가결과별로 구분해서 보았을때도 신의료기술, 연구단계기술 각각 근거의 수준을 미수행한 경우가 64%, 63%로 가장 높은 비율을 차지하였다. ⅲ) 영상검사에서는 전체적으로 근거의 수준 C가 56%로 가장 높은 비율을 차지하였으며 평가결과별로 구분해서 보았을때는 신의료기술은 근거의 수준 C가 63%, 연구단계기술의 경우 근거의 수준 D가 58%로 가장 높은 비율로 보고하였다. ⅳ) 기타 검사에서는 전체적으로 근거의 수준 C가 51%로 가장 높은 비율로 보고되었으며 평가결과별로 구분해서 보았을때는 신의료기술은 근거의 수준 C 53%, 연구단계기술은 근거의 수준 D가 48%로 가장 높은 비율로 보고되었다.

표 43. 처치 및 시술에서 평가결과별 근거의 수준

단위: 건(%)

처치 및 시술	근거의 수준					
	합계	A	B	C	D	미수행
신의료기술	192(100)	13(7)	67(35)	29(15)	76(40)	7(4)
연구단계기술	160(100)	7(4)	22(14)	16(10)	93(58)	22(14)
계	353(100)	20(6)	89(25)	45(13)	169(48)	29(8)

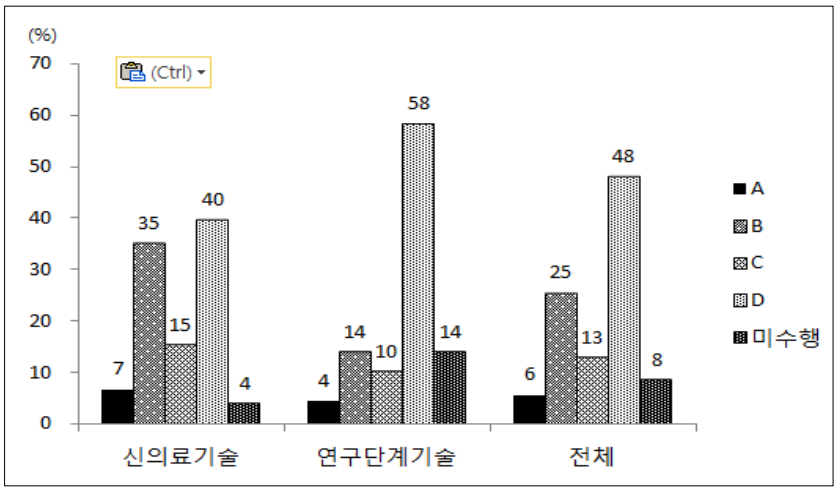

그림 13. 처치 및 시술에서 평가결과별 근거의 수준

표 44. 체외진단검사에서 평가결과별 근거의 수준

단위: 건(%)

체외진단	근거의 수준				미수행	합계
	A	B	C	D		
신의료기술	1(0.2)	5(1)	123(24)	55(11)	324(64)	508(100)
연구단계기술	1(0.4)	3(1)	47(18)	49(18)	167(63)	267(100)
계	2(0.3)	8(1)	170(22)	104(13)	491(63)	775(100)

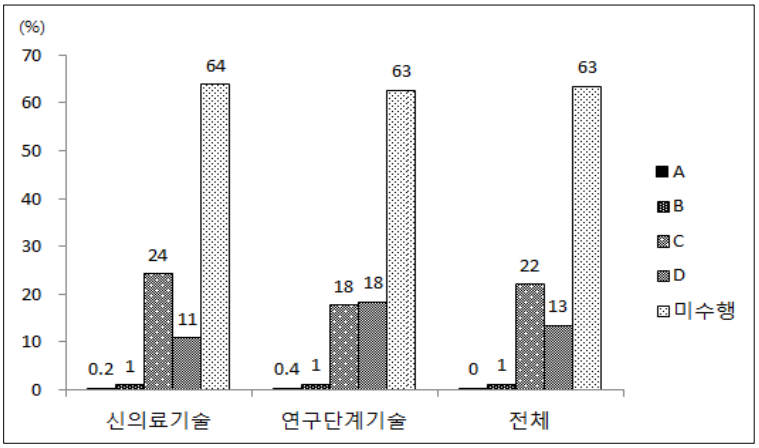

그림 14. 체외진단검사에서 평가결과별 근거의 수준

표 45. 영상검사에서 평가결과별 근거의 수준

단위: 건(%)

영상검사	근거의 수준				미수행	합계
	A	B	C	D		
신의료기술	0	4(11)	22(63)	7(20)	2(6)	35(100)
연구단계	0	0	8(42)	11(58)	0	19(100)
계	0	4(7)	30(56)	18(33)	2(4)	54(100)

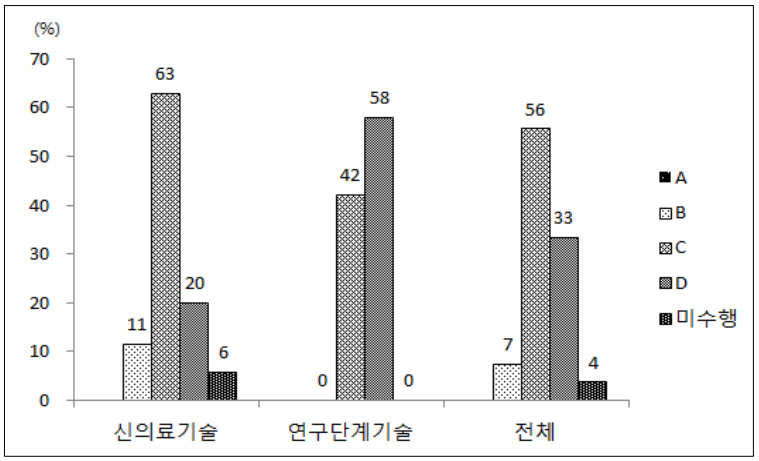

그림 15. 영상검사에서 평가결과별 근거의 수준

표 46. 기타검사에서 평가결과별 근거의 수준

단위: 건(%)

기타검사	근거의 수준				미수행	합계
	A	B	C	D		
신의료기술	0	3(5)	34(52)	20(31)	8(12)	65(100)
연구단계기술	0	0	15(45)	16(48)	2(6)	33(100)
계	0	3(3)	49(50)	36(37)	10(10)	98(100)

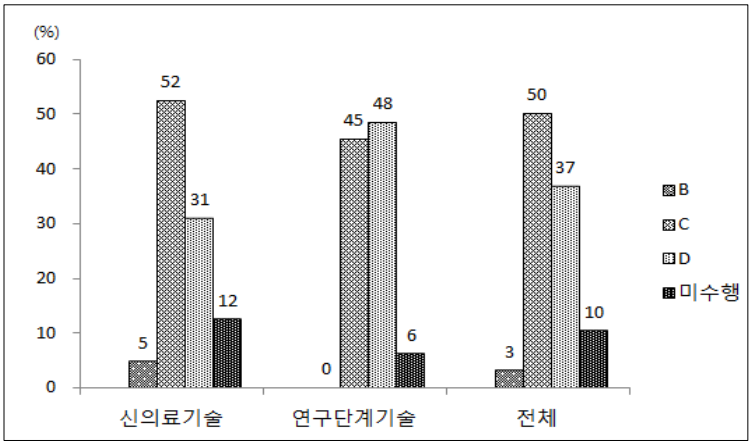

그림 16. 영상검사에서 평가결과별 근거의 수준

다, 신의료기술로 인정된 기술의 근거수준에 따른 급여 유형 비율 분석

신의료기술 평가가 수행되어 신의료기술로 인정된 기술 중 급여권에 등재된 기술에 대한 검토를 수행하였다. 이미 언급한 바와 같이 신의료기술 평가 신청 후 평가가 수행된 기술은 총 1,279건이었으며, 이 중 신의료기술로 인정된 기술은 총 800건이었고, 급여권에 등재된 기술은 총 403건으로 확인되었다.

급여 등재 기술의 근거 수준 평가결과를 검토한 결과, 이 경우에도 신속심의 등으로 근거 수준 평가 없이 신의료기술 기술로 인정되어 근거 수준 평가가 미수행된 것이 180건(45%)으로 가장 높은 비율을 차지하고 있었다. 근거 수준 평가가 수행된 기술들 중에서는 근거의 수준 C로 평가된 기술이 106건(26%)으로 많았으며, 그 다음으로는 근거의 수준 D가 69건(17%), 근거의 수준 B가 40건(10%), 근거의 수준 A가 8건(2)% 순이었다.

급여 등재 기술의 근거 수준 평가 결과를 급여 유형에 따라 구분하여 분석한 결

과에서는 급여로 등재된 기술과 선별급여로 등재된 기술의 근거의 수준 카테고리별 순위에 약간의 차이가 있는 것을 확인할 수 있었다. 검토 결과, 급여로 등재된 기술은 전체 분석 결과와 동일하게 근거 수준 평가가 미수행된 기술의 비율이 48%로 가장 높았고, 그 다음이 근거의 수준 C 24%, D 17%, B 9%, A 2% 순이었으나, 선별급여로 등재된 기술에서는 근거의 수준 C로 평가된 기술의 비율이 46%로 가장 높고, 근거의 수준 B와 D로 평가된 기술이 각각 22%, 근거 수준 평가가 미수행된 기술이 11%로 확인되었으며, 근거의 수준 A로 평가된 기술은 없는 것으로 나타났다.

신의료기술로 인정되었으나 아직 급여권에 등재되지 않은 비급여 기술에서는 급여 등재기술의 전체 분석 결과와 동일하게 근거 수준 미수행 기술이 41%로 가장 높았고, 이 후 근거의 수준 C 25%, 근거의 수준 D 24%, 근거의 수준 B 9%, 근거의 수준 A 1% 순이었다.

표 47. 신의료기술로 인정된 기술의 급여 유형별 근거의 수준

단위: 건(%)

근거수준	급여 등재 기술			비급여	합계
	급여	선별급여	합계		
A	8(2)	0	8(2)	3(1)	11(2)
B	32(9)	8(22)	40(10)	18(9)	58(9)
C	89(24)	17(46)	106(26)	52(25)	158(26)
D	61(17)	8(22)	69(17)	49(24)	118(19)
미수행	176(48)	4(11)	180(45)	86(41)	266(44)
계	366(100)	37(100)	403(100)	208(100)	611(100)

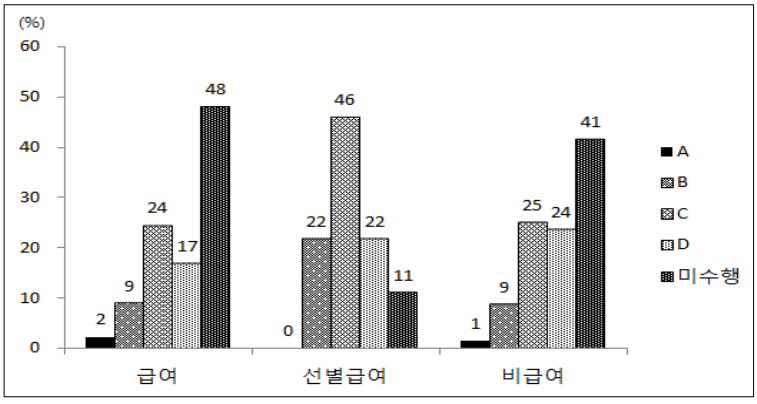

그림 17. 신의료기술로 인정된 기술의 급여 유형별 근거의 수준

라, 의료행위전문평가위원회의 급여 결정 판단 근거 분석

신의료기술로 인정된 기술 중 건강보험심사평가원 홈페이지 게시자료를 통해 의료행위전문평가위원회의 급여 결정시 판단 근거를 공개한 기술 127건에 대해 의사결정 판단 근거에 대한 현황을 분석하였다.

먼저 공개된 자료의 127건을 급여 유형별로 구분해본 결과 급여로 결정된 기술이 80건(63%)로 가장 많았으며, 비급여가 27건(21%)로 두 번째, 선별급여가 20건(16%)으로 가장 적었다.

표 48. 신의료기술로 인정된 기술 중 급여 결정 판단 근거가 공개된 기술의 급여 현황

구분	건수	%
급여	80	63
선별급여	20	16
비급여	27	21
합계	127	100

의사결정 판단 근거는 공개된 자료들을 모두 검토한 후 비슷한 유형의 표현끼리 통합하는 과정을 통해 다음의 10가지 항목으로 정리하였다.

1) 치료효과성이 우수한 기술
2) 치료효과성이 기존기술과 유사한 기술
3) 치료방향 결정에 도움을 줄 수 있는 기술
4) 기존 검사의 보조검사, 추가검사로 사용되거나 기존 치료의 보완적 시술로 사용될 수 있는 기술
5) 문헌적 근거수준의 고려 또는, 교과서 및 가이드라인, 임상진료지침에 사용이 권고되어있거나 관련 언급이 있는 기술
6) 신속성을 높여주는 기술
7) 편의성이 있는 기술
8) 덜 침습적인 기술
9) 사회적 요구도가 높은 기술
10) 대체가능한 다른 치료방법이 없는 기술

위의 10가지 항목을 기준으로 급여 유형별 의사결정 판단 근거의 언급 빈도를 분석한 결과, 급여로 인정된 기술들 중 의사결정 시 가장 많이 언급된 근거 요소는 기술의 임상적 유효성과 관련된 요소들로, 치료효과성이 우수하거나 기존기술과 유사한 정도 혹은 치료방향 결정에 도움을 주는 기술로 받아들여지고 있는 기술인 경우가 가장 높은 비율을 차지하고 있었다. 선별급여로 결정된 기술들에서는 편의성이 있거나 덜 침습적인 기술, 사회적 요구도(유병률이 높은 환자를 대상으로 하거나 법정전염병과 같이 빠른 정책적 결정이 필요한 기술)가 높은 기술이 판단 근거의 요소로서 많이 언급되고 있었다. 비급여로 결정된 기술들에서는 이미 기존 검사나 기존 시술적 치료방법이 존재하여 신청 기술이 사용목적이 보조검사나 추가검사, 또는 보완적 시술로 예상되는 경우에 대한 언급이 가장 빈번하게 제시되어 있었다.

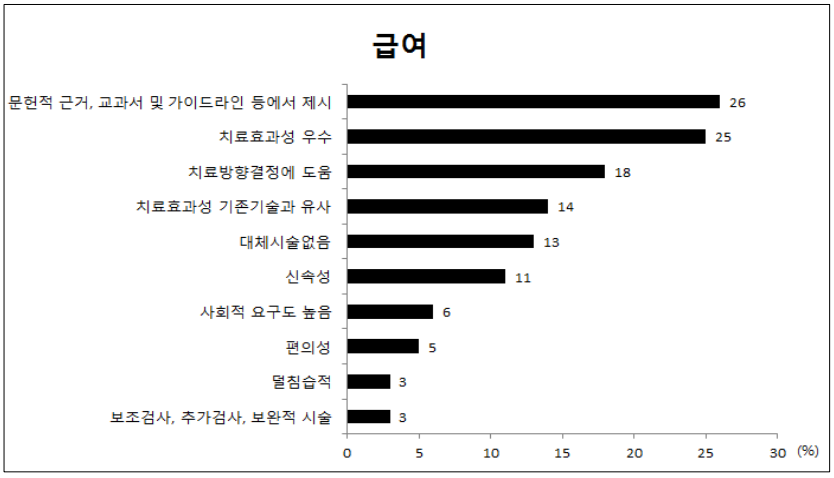

그림 18. 급여기술에서 의사결정 판단 요소 비율 비교

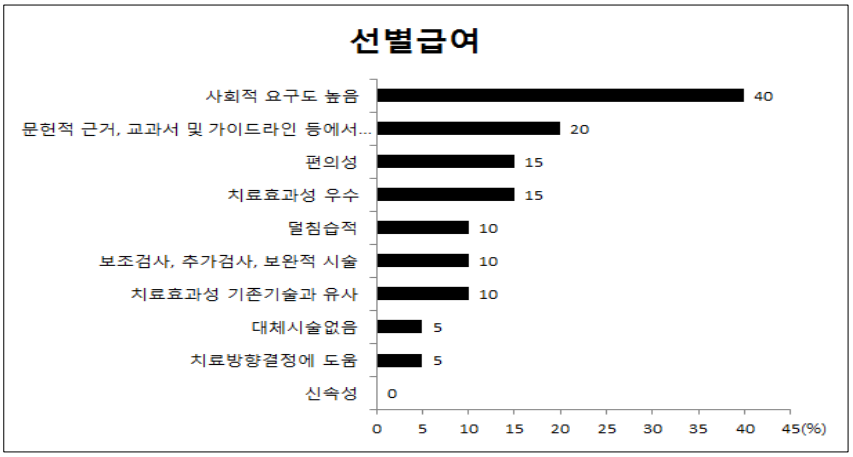

그림 19. 선별급여기술에서 의사결정 판단 요소 비율 비교

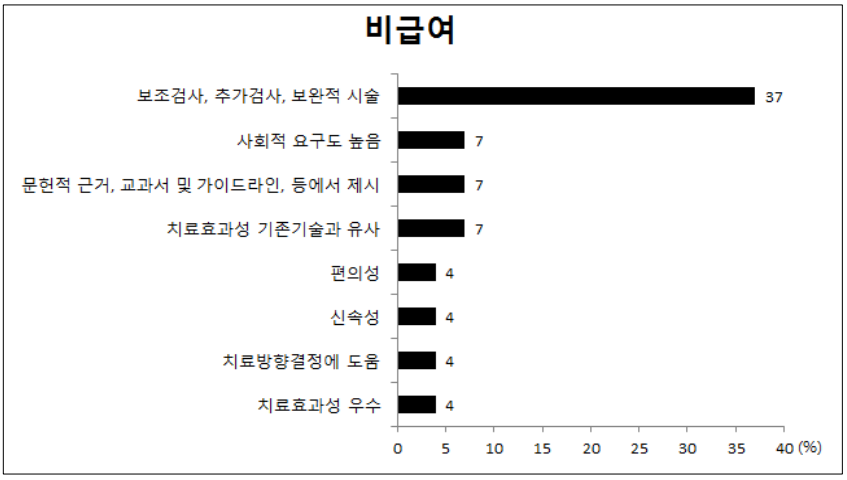

그림 20. 비급여기술에서 의사결정 판단 요소 비율 비교

기술 유형별 근거의 수준에 따른 급여 결정 시 중요하게 논의되었던 사항들의 자세한 내용은 다음의 표 49 ~ 표 52와 같다.

표 49. (처치 및 시술)급여유형별 급여 결정 시 의결사항 특징

근거 수준	급여 유형	급여 결정 시 의결사항의 특징
A	급여	– 질병중증도 있고, 고가 장비사용 – 치료효과성(임상적 유용성) 있음
	비급여	– 임상적 유용성이 확인되지 않고 비용효과성이 불분명
B	급여	– 치료효과성 있음 – 기존기술에서 적응증 확대
	선별급여	– 치료효과성(임상적 유용성) 있으나 남용의 우려가 있으며 안전사고를 위해 적절한 규모의 의료기관에서 시행 필요 – 치료효과성 있으나 비용효과성 입증 필요
	비급여	– 비용효과성이 불분명함 – 치료효과성 있으나 추가 근거가 필요하고 비용효과성이 낮음 – 기존의 유사기술이 비급여로 등재 – 보완적으로 실시하는 기술이며 비용효과성이 낮음 – 기존 비급여로 등재된 유사기술과 치료효과성은 유사, 비용효과성 불분명
	비급여→선별급여	– 치료효과성 부족하고 국내 근거자료 축적이 필요하며 비용효과성 낮음
C	급여	– 기존의 급여 등재기술에서 대상 적응증 확대
	선별급여	– 비용효과성 불분명하나, 사회적 요구도 고려
	비급여 → 선별급여	– 치료효과성(임상적 유용성)있으나 고가이고 사회적 요구도 낮음
D	급여	– 치료효과성(임상적 유용성) 있고 대체가능한 기술 없음 – 기존 유사기술이 급여로 등재되어 있으며 대체가능한 기술이 없음 – 고가의 기술, 고난이도 기술로 중점관리 필요하나 치료효과성 있고, 대체가능한 기술이 없음 – 질병중증도, 치료효과성 있음 – 치료효과성 있고 기존기술에 비해 비용이 저렴함
	선별급여	– 비용효과성 불분명하나 편의성 향상되고 사회적 요구도 있음 – 대체행위가 없는 기술이나 임상적 유용성에 대한 근거 축적 필요
	비급여	– 임상근거 부족 – 치료효과성 낮고 장기추적관찰결과가 필요, 비용효과성 낮음
R	급여	– 치료효과성(임상적 유용성) 있고 대체기술 없음
	선별급여	– 치료효과성(임상적 유용성) 있으나 고비용

표 50. (체외진단검사)급여유형별 급여 결정 시 의결사항 특징

근거의 수준	급여 유형	급여 결정 시 의결사항의 특징
B	급여	– 교과서, 진료지침에서 임상적 유용성 제시되었고 대체가능한 검사가 없음
C	급여	– 치료효과성(임상적 유용성) 있음 – 치료효과성(임상적 유용성) 있으며, 비용 저렴, 또는 비용효과성이 있음 – 치료효과성(임상적 유용성)이 있고 대체기술 없음 – 고가의 재료가 소요되나 치료효과성 있고 기존행위 일부 대체 가능. 다른 적응증에서는 급여로 등재됨
	선별급여	– 치료효과성이 예상되나 임상적 유용성에 대한 자료 축적이 필요 – 기존검사의 한계를 보완가능하나 고가재료가 사용됨. 추가근거도 필요함 – 비용효과성 불분명하나 사회적 요구도 고려
	비급여	– 보완검사, 치료효과성, 비용효과성 불분명, 임상적 근거 부족 – 추가검사이며, 신뢰도가 높지 않고 임상적 유용성에 대한 근거 축적 필요 – 비용효과성 낮고 임상적 유용성 낮음
	비급여→급여	– 보조검사이나 해당질환 진단의 참고표준이 없고 임상적 유용성 입증됨
	선별급여→급여	– 기존기술과 상호보완적으로 사용시 비용효과성 있음 – 치료효과성 있음
D	급여	– 치료효과성 있음 – 기존 기술이 급여기술이면서 사회적 요구도 있음
	선별급여	– 치료효과성(임상적 유용성) 있고, 비용효과성이 불분명하나, 사회적 요구도가 높음
	비급여	– 고가의 재료 소요되며 치료효과성이 불확실함 – 예후 예측 등 목적의 유전자검사의 경우 타 질환에서 비급여 적용한 점과 급여 정책을 고려함
R	급여	– 비용효과성이 낮고 치료효과성이 우수함 – 치료효과성 우수 – 고가의 장비이나 치료효과성(임상적 유용성)이 있음 – 치료효과성 우수, 유전자 검사 급여전환원칙 고려 – 유사기술이 급여로 등재되어있고 해당균주의 임상적 유효성 확인 – 사회적, 정책적 요구도가 있음(법정전염병) – 임상적 유용성이 있으며 대체 행위가 없음
	선별급여	– 임상적 유용성이 있으나 비용효과성 불분명. 추가 근거축적이 필요함 – 치료효과성(임상적 유용성) 있으나 임계값 정해지지 않음
	비급여	– 치료효과성 및 비용효과성이 불분명함 – 대체검사가 없고 임상적 유용성 입증되었으나 해당질환이 타법령에 의해 보상되고 건강검진의 목적으로 사용되고 있음

근거의 수준	급여 유형	급여 결정 시 의결사항의 특징
		– 치료방향 결정에 미치는 영향이 없음 – 사회적 요구도가 있으나 치료효과성 및 비용효과성이 불분명함 – 기존기술과 진단정확성이 유사한 수준이나 근거문헌이 부족하고 임상적 유용성에 대한 논란 있으며 고가의 시약을 사용해야함
	비급여→급여	– 정책적 요구도 있음 – 기존검사의 일부를 대체 가능하고 비용효과성이 유사함

표 51 (영상검사)급여유형별 급여 결정 시 의결사항 특징

근거의 수준	급여 유형	급여 결정 시 의결사항의 특징
B	급여	– 치료효과성이 우수함 – 임상적 유용성이 있고 대체가능한 검사가 없음
C	급여	– 대상 질환이 많지 않고 기존 검사를 일부 대체 가능하며 치료효과성 우수하고 비용효과성이 인정됨 – 치료효과성 있음
	비급여	– 치료효과성 없음. 보조검사 목적의 검사이고 비용효과성이 낮음 – 치료효과성이 낮고 임상적 유용성에 대한 논란이 있음
R	비급여	– 임상적 유용성이 불분명함. 비용효과성이 낮고 유사검사가 비급여임

표 52 (기타검사)급여유형별 급여 결정 시 의결사항 특징

근거의 수준	급여 유형	급여 결정 시 의결사항의 특징
C	급여	– 치료효과성이 우수하고 비침습적인 기술이며 비용효과성이 있음 – 치료효과성이 우수함
	선별급여	– 기존기술에 비해 비침습적이고 응급상황에 유용하나 치료효과성이 낮고 비용효과성이 불분명함
	비급여	– 치료효과성 및 비용효과성이 불분명함 – 임상적 유용성을 판단할만한 자료가 부족하고 치료효과성이 낮으며 비용효과성도 불분명함
D	급여	– 기존기술과 치료효과성 및 비용효과성이 유사함
	선별급여	– 임상적 근거자료의 축적이 필요함. 비용효과성이 불분명하고 고가의 치료재료가 소요됨. 임상적 유용성은 교과서에서 확인 가능함
	비급여	– 질환의 특성상 여러 급여 기존검사들과 환자의 증상을 종합적으로 고려하여 진단하고 있는 상태이며 해당 검사는 아직까지 임상에 도입된 기간이 짧아 보편적으로 사용하고 있지 않음 – 편의성, 신속성이 있으나 진단을 위해 추가검사가 필요하고 유사행위가 비급여로 등재되어있으며 임상적 유용성 확인을 위한 자료 축적이 필요함
R	급여	– 교과서 및 다수문헌에서 확립된 검사로 제시되어 있음 – 기존 유사기술이 급여로 등재되어 있으며, 기존검사의 보완검사임. 간단한 검사이므로 진찰료, 입원료 소정점수에 포함시키도록 함

근거의 수준	급여 유형	급여 결정 시 의결사항의 특징
		- 치료효과성(임상적 유용성) 있고 대체가능한 기술이 없음
	선별급여	- 기존검사도 임상에서 거의 시행되지 않고 유사기술이 선별급여로 등재됨

급여 유형별로 급여 결정시 중요하게 고려되었던 판단 근거를 분석해본 종합적인 결과는 다음과 같다.

ⅰ) 급여기술의 경우
① 질병의 중증도가 있고,
② 치료효과성(임상적 유용성)이 높거나 입증된 기술,
③ 대체가능한 기술이 없는 경우,
④ 사회적(정책적) 요구도가 높은 기술,
⑤ 비용효과성 있는 경우,
⑥ 기 결정사항을 고려하여 의사결정이 된 기술인 경우 급여 기술로 결정되는 경향이 있는 것을 확인할 수 있었다.

ⅱ) 선별급여기술의 경우
① 치료효과성(임상적 유용성)이 예상되나 근거자료축적 필요,
② 사회적 요구도 높은 기술,
③ 비용효과성 불확실한 경우,
④ 편의성이 증대된 기술,
⑤ 남용의 우려가 있는 기술인 경우 선별급여 기술로 결정되는 경향이 있는 것을 확인할 수 있었다.

ⅲ) 비급여 기술의 경우
① 치료효과성(임상적 유용성) 부족하거나 불분명한 기술,
② 장기추적관찰 필요,
③ 비용효과성 불분명한 기술(고가 기술 등),
④ 기 결정 사항 고려하여 의사결정 된 기술인 경우 비급여 기술로 결정되는 경향이 있는 것을 확인 할 수 있었다.

마, 신의료기술평가의 근거 수준 평가와 급여 결정 판단에 대한 사례 비교분석

이전까지의 현황 분석 결과를 검토하였을 때, 신의료기술로 인정된 기술들 중 많은 사례들에서 신의료기술평가의 근거수준 평가결과와 급여 결정의 결과 사이에 갭(gap)이 존재하는 것이 확인되어 이에 대한 이유를 탐색하기 위하여 대표 사례들을 중심으로 각 평가에서 논의되었던 검토사항에 대해 비교분석을 실시하였다.

신의료기술 평가 시 근거의 수준 A로 신의료기술로 인정된 기술 중 비급여 행위로 결정된 기술로는 만성두드러기 환자에서 자가혈액 또는 자가혈청 근육주사요법이 확인되었다. 동 기술의 신의료기술평가는 무작위배정임상연구 문헌 3편과 증례연구 4편을 바탕으로 수행되었으며, 문헌들에서 시술 관련 심각한 합병증 발생이 보고된 바 없고 환자의 증상과 약물 사용량 등의 유효성 지표에서 중재시술로 인한 개선 효과가 확인되었다. 이러한 결과를 바탕으로 동 기술은 증상이 조절되지 않는 만성 환자에서 대체기술이 없고 치료에 대한 환자의 요구도가 있는 기술로 인정받아 신의료기술로 인정되었으며 무작위배정임상시험 연구 3편의 근거를 바탕으로 근거의 수준은 A로 평가되었다. 하지만 이 후 급여 여부 평가에서 동 기술은 국내·외 교과서 등에서 임상적 유용성에 대한 언급이 없고 비용효과성이 불분명하며 신의료기술 인정 후 국내 실시현황이 미미하여 비급여 기술로 결정되었다. 평가 결과의 갭과 관련하여 살펴보면, 임상적 유용성에 대한 근거 자료 확인에 있어 신의료기술평가에서는 체계적 문헌고찰을 통한 일차연구들에서 해당 기술의 안전성 및 유효성을 확인하는 방법으로 평가를 수행하였지만 급여 결정 시에는 국내·외 교과서 등의 임상적 유용성에 대한 언급 여부를 검토하여 이 부분에 있어 검토한 근거의 종류가 달랐던 점을 확인할 수 있었다. 또한 급여 결정의 경우 공공의 재정이 소요되어야 하기 때문에 신의료기술평가에서는 고려하지 않는 비용 효과성에 대한 항목을 중요하게 고려하고 있는 것을 확인 할 수 있었다.

표 53 근거의 수준이 A이면서 비급여로 결정된 기술의 사례분석: 자가혈액 또는 자가혈청 근육주사요법(만성 두드러기 환자)

신의료기술평가 결과(근거의 수준 A)	급여 결정사항(비급여)
〈검토문헌〉 - 무작위배정임상연구 3편(1++(1편), 1-(2편)), 증례연구 4편 〈안전성〉 - 시술관련 심각한 합병증 발생 없음 〈유효성〉 - 증상 개선 정도, 약물 사용량 변화 및 삶의 질이 개선되는 결과 〈결론〉 - 표준적인 약물치료로 증상이 조절되지 않는 만성환자로, 대체가능기술 부재하고, 치료에 대한 환자 요구도 높아 임상적 의의 있음	- 국내·외 교과서 등에서 임상적 유용성에 대해 언급하고 있는 것이 없음 - 비용효과성이 불분명함. 기존치료인 약물투여, 탈감작요법 등을 수행 후 이 기술을 위해 추가적 비용을 지불해야함 - 국내 1개 기관만 실시하고 있어 실시현황이 미미함

신의료기술 평가 시 근거의 수준 D로 신의료기술로 인정된 기술 중 급여 행위로 결정된 기술로는 양막부착술이 확인되었다. 동 기술의 신의료기술평가는 코호트 연구 2편과 증례연구 9편을 바탕으로 수행되었으며, 평가 시 비교연구 문헌이 충분치는 않았으나 질환의 특성상 비교연구를 수행하기 어려운 점과 대체기술이 없는 점을 고려하면 현재 제시된 결과들에서 안전성 및 유효성이 수용가능한 수준이라 평가되어 신의료기술로 인정되었고 근거의 수준은 가장 낮은 등급인 D로 결정되었다. 이 후 급여 여부 평가에서 동 기술은 조산아 출생률이 증가하는 현재의 사회적 상황을 고려하였을 때 사회적 요구도가 있는 질환 관련 기술이며, 대상 환자의 생존율에 영향을 미칠 수 있고, 대체가능한 기존 기술이 없는 점 등을 고려하여 급여 기술로 결정되었다. 동 사례에서는 신의료기술평가와 급여 결정 모두에서 대체가능한 기술이 없다는 점이 공통적으로 언급되면서 각각의 평가에서 모두 긍정적인 방향으로 결과가 결정된 것을 확인 할 수 있었다.

표 54 근거의 수준이 D이면서 급여로 결정된 기술의 사례분석: 양막부착술

신의료기술평가 결과(근거의 수준 D)	급여 결정사항(급여)
〈검토문헌〉 - 코호트2편(2+, 2-), 증례연구 9편 〈안전성〉 - 시술관련 심각한 합병증 발생이 없고, 양막파열이 모체와 태아 감염 및 산과적 합병증 위험에 노출되어 있어 안전성 수용가능 〈유효성〉 - 대조군 비교 시 유의한 차이 없으나, 사례들을 종합한 결과에서 임신유지 기간 연장, 시술성공률, 주산기 및 신생아 예후 평가결과 수용가능 〈결론〉 - 비교 연구 결과가 부족하나, (i) 질환의 특성상 비교연구를 수행하기 어려운 상황이며, (ii) 시급한 치료를 요하는 고위험 질환임에도 불구하고 대체가능 기술이 없다는 점을 감안 시 유효성 결과 수용가능	- 시술빈도 낮으나 조산아 출생률 증가하면서 사회적 요구도 높음 - 태아의 재태연령에 따라 생존율에 영향을 미침 - 대체가능한 기술 없음

신의료기술로 인정되었으나 체계적 문헌고찰 방법이 아닌 신속검토로 평가가 수행되어 근거 수준 평가가 미수행된 기술들 중 급여로 결정된 기술과 비급여로 결정된 기술을 확인하여 각각 1건의 사례에 대해 각 평가에서 논의되었던 검토사항에 대해 비교분석을 실시하였다.

신의료기술평가 시 신속검토로 근거수준이 평가되지 않았으나 급여로 결정된 사례는 CMV 약제내성 유발 돌연변이[염기서열분석] 검사가 확인되었다. 동 기술의 신의료기술평가는 체계적 문헌고찰에 대한 수행 없이 증례연구 4편과 증례보고 3편의 문헌을 신속으로 검토하는 형태로 수행되었으며, 평가 결과 약제 내성을 확인하고 치료약제를 결정하는데 임상적 유용성이 있으며 이미 교과서에도 기술된 확립된 검사방법임이 확인되어 신의료기술로 신속 인정되었다. 이 후 급여 여부 평가에서 동 기술은 이미 신의료기술로 인정을 받은 안전하고 유효한 검사이며 국내·외 가이드라인에서 검사의 실시를 권고하고 있고 투여 약제에 대한 내성을 파악하여 투여 약제의 변경 등에 영향을 줄 수 있는 검사임을 고려하여 급여 기술로 결정되었다. 동

기술의 경우 신의료기술평가를 통해 내려진 결론이 급여 결정 시 고려되었던 사항으로 언급되어 신의료기술평가의 결과가 급여 결정에도 영향을 미쳤음으로 확인 할 수 있는 사례였다.

표 55. 근거 수준 평가를 수행하지 않은 기술이면서 급여로 결정된 기술의 사례분석: CMV 약제내성 유발 돌연변이[염기서열분석]

신의료기술평가 결과 (근거수준평가 미수행)	급여 결정사항(급여)
〈신속검토(체계적 문헌고찰 미수행)〉 – 증례연구 4편, 증례보고 3편 〈결론〉 – 약제내성을 확인하고 치료약제를 결정하는데 임상적 유용성이 있음 – 염기서열검사는 배양검사의 대체적 참고표준검사로 사용되고 있고, 교과서에도 기술된 확립된 검사방법임 – 다만, 사용대상과 관련하여 거대세포바이러스 감염환자 중 항바이러스 치료를 지속적으로 받음에도 불구하고 거대세포바이러스 양이 2주 이상 지속되거나 증가하는 환가로 제한하는 것이 적절함	– 신의료기술 평가 결과 안전하고 유효한 검사로 평가되었음 – 국내·외 가이드라인에서 항바이러스제 약제내성 여부 확인을 위해 동 검사의 실시를 언급하거나 권고하고 있음 – 투여 약제에 대한 내성 여부를 파악하여 내성 돌연변이가 있을 경우 투여 약제의 변경 혹은 투여 용량 변화 등 치료에 도움을 줄 수 있음

신의료기술평가 시 신속검토로 근거수준이 평가되지 않았으나 비급여로 결정된 사례는 인플루엔자바이러스 RNA 검사[등온증폭-교잡반응법] 검사가 확인되었다. 이전 사례와 마찬가지로 동 기술 역시 신의료기술평가는 체계적 문헌고찰에 대한 수행 없이 진단법평가연구 4편을 포함한 관련문헌 6편을 신속으로 검토하는 형태로 수행되었으며, 평가 결과 임상적 유용성 및 진단적 타당성이 있는 기술로 평가되어 신의료기술로 신속 인정되었다. 하지만 이 후 급여 여부 평가에서 동 기술은 검사가 질환의 치료에 미치는 영향과 비용효과성이 불분명한 점을 고려하여 비급여 기술로 결정되었다. 각각의 평가 결과 간의 갭을 살펴보면, 신의료기술평가의 경우 관련 문헌의 검토를 통해 동 검사의 임상적 유용성과 진단적 타당도를 검토하여 신의료기술로 인정하였으나, 급여 여부 결정에서는 검사 자체의 진단적 타당도에 대한 언급 없이 검사 결과가 치료에 미치는 영향력과 건강보험재정에 미치는 비용을

고려하여 검사의 경제성에 대한 항목을 의사 결정 시 주요 요소로 고려한 것을 확인 할 수 있었다.

표 56. 근거 수준 평가를 수행하지 않은 기술이면서 급여로 결정된 기술의 사례분석: 인플루엔자바이러스 RNA 검사[등온증폭-교잡반응법]

신의료기술평가 결과 (근거수준평가 미수행)	급여 결정사항(비급여)
〈신속검토(체계적 문헌고찰 미수행)〉 – 관련문헌 6편(진단법평가연구 4편 포함) 〈유효성〉 – 기존기술과 사용목적 및 대상은 유사하나 검사방법이 다름 – 인플루엔자 A와 B 바이러스를 검출하는 검사로 임상적 유용성 및 진단적 타당도 있음	– 인플루엔자감염 진단에 대한 사회적 요구도는 있음 – 인플루엔자 바이러스에 대한 특이적인 치료법이 없어 검사가 치료에 미치는 영향이 불분명함 – 고가의 시약을 사용, 대상자가 많은 점 등 건강보험 재정에 미치는 영향이 큰 점을 감안하면 현 시점에서는 진료상의 경제성이 불분명함

이러한 사례분석 결과를 종합하였을 때, 신의료기술 평가와 급여 결정 평가 시 고려된 항목과 자료가 대상 기술에 따라 유사한 경우와 상이한 경우가 모두 존재하는 것이 확인되었으며, 각각의 평가 모두 결국은 확인 가능한 근거를 바탕으로 전문가 집단의 종합적인 합의의 과정을 통해 결론이 내려지고 있음을 확인 할 수 있었다.

4. 신의료기술평가 권고체계 개선방안

4.1. 개요

신의료기술평가에서 권고등급을 결정하기 위한 목적은 신의료기술평가 결과가 시장진입된 이후 임상현장에서의 효율적 사용을 도모하고, 급여결정과정과 의료기술평가 전주기적 관점에서 재평가 대상선별을 위한 합리적인 의사결정을 지원할 수 있도록 연계 방안을 마련하는데 있다. 본 연구에서는 현재 수행 중에 있는 신의료기술평가의 세부절차 중 결론을 도출하는 과정에서 의과학적인 근거를 기반으로 하여 다양한 요인들을 고려한 가치를 판단하여 합리적인 의사결정을 돕는 권고체계를 제안하고자 하였다.

신의료기술의 권고체계를 정하고, 등급을 판단하는 기준에 대해서는 신의료기술평가가 시장진입된 후 의사결정되었던 사례를 토대로 분석한 내용과 국내외 의료기술평가기관 및 임상진료지침 등을 참고하여 초안을 마련하였으며, 이후 전문가 자문과 권고시범 적용을 통해 현 신의료기술평가 절차 내 실용적인 측면에서 제한점 및 보완이 필요한 사항 등을 검토하였다.

4.2. 권고체계(안)

4.2.1. 권고체계(등급)

권고체계는 현재의 신의료기술평가제도의 결정구조를 벗어나지 않는 범위내에서, 신의료기술(권고함), 혁신의료기술(조건부권고), 연구단계기술(권고하지 않음)의 3등급 체계로 분류하였고, 신의료기술에 대한 권고등급은 ‘가 - 권고강도 강함’과 ‘나 - 권고강도 약함’으로 제시하였다. 혁신의료기술의 조건부 권고에 대해서는 문헌적 근거로써는 유효성의 불확실성이 있는 기술 중 잠재적 가치를 인정하여 재평가를 조건으로 권고한다는 점을 고려하여, 권고의 강도(등급) 체계에서는 별도로 제시하지 않았다. 그 외, 신의료기술평가 유예에 해당되는 기술도 신의료기술의 평가를 받지 않았기 때문에 별도로 권고등급을 판단하는데 고려대상에서 제외하였다.

신의료기술에 대한 권고등급은 3등급 체계(권고강도 강함, 중등도, 약함)도 고려

하였으나 중등도(moderate)에 대한 기준이 애매하고 척도화하지 않는 합의결정구조에서는 2등급 체계가 보다 쉽게 의사결정하는데 용이할 수 있다는 전문가 의견을 반영하여 2등급 체계로 구성하였다.

또한, 신의료기술에 대한 의료기술재평가를 연속성 있게 수행하기 위해서 재평가를 조건으로 인정한 '혁신의료기술'과 '권고강도가 낮은 신의료기술'에 대해 재평가가 우선적으로 필요한 기술로써 제안하였다. 그러나 신의료기술은 임상에서 사용되기 전의 제한된 근거로 도입된 것으로써 시장도입된 이후에 의료기술에 대한 근거와 가치판단은 사회적 환경 및 평가시점에 따라 달라질 수 있으므로 과거의 평가척도만으로 의료기술재평가가 수행되어서는 아니되고, 재평가를 수행하는 단계에서 우선순위에 대한 선별기준을 토대로 재평가 기술을 선정하는 것이 보다 바람직할 것으로 보인다.

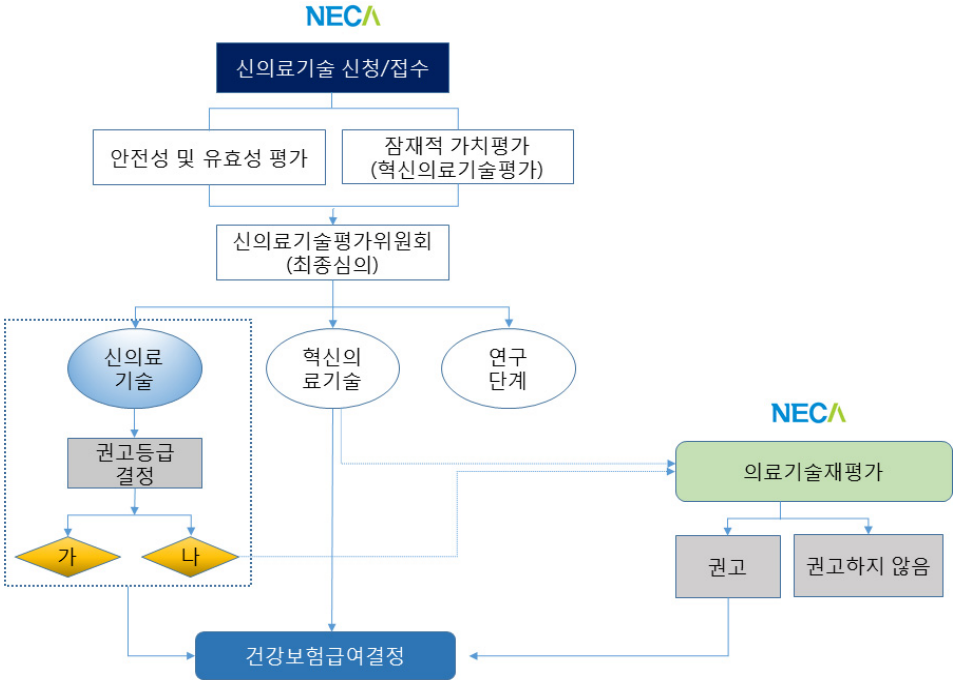

그림 21. 신의료기술평가 권고체계(안)

표 57. 권고체계(안)

권고체계		설명
구분	등급*	
신의료기술	가	권고강도 강함
	나	권고강도 약함
혁신의료기술	조건부 권고	
연구단계기술	NA	

* 권고등급은 권고 평가항목을 종합적으로 고려하여 등급을 분류함
NA, 연구단계기술은 임상에서 사용이 불가한 기술로써 '권고하지 않음'으로 제시하는 것이 오히려 임상사용의 권고정도와 혼동될 수 있어, 본 연구에서는 제외되었음

4.2.2. 권고등급 평가기준

신의료기술의 권고등급 체계는 권고를 누구에 겨냥하여 내리는지, 쓰이는 관점에 따라 평가기준이 달라질 수 있어 이러한 점을 고려하여 2가지 평가기준 방안을 마련하였다. 즉, 급여결정을 지원하기 위한 권고로써 사용할 것인지, 신의료기술평가 보고서의 문헌적 근거자료를 토대로 임상적 측면에서 가치 평가를 보완하여 권고강도로 사용할 것인지 그 목적과 활용성에 따라 평가기준을 다르게 접근하였다.

가, 평가기준(1안)

신의료기술로 인정 받은 기술은 건강보험 등재시 활용된다는 점을 고려하여 보험등재수준 결정과의 연계 강화 차원의 정부 및 보험자적인 관점에서 접근하였다.

신의료기술로 인정되어 시장진입된 기술에 대한 요양급여 의사결정은 '요양급여 대상의 여부 결정에 관한 규칙' 에 따라, 의학적 타당성, 의료적 중대성, 치료효과성, 비용효과성, 환자의 비용부담 정도 및 사회적 편익 등을 고려하여 숙의과정을 통해 수행되어지고 있다. 신의료기술의 임상도입 과정은 신의료기술평가와 요양급여결정의 단계로 이분화되어 있다는 점에서 두 제도의 목적과 취지가 상이하고 평가의 판단기준이 다르기 때문에 신의료기술평가 결과를 급여 결정시에 직접적으로 원용하여 활용되기에는 제한점이 있으나, 보험권 내에서 의사결정과정시에 신의료기술평가 근거자료의 활용을 도모하여 임상적 유용성 판단을 위한 의사결정에 도움을 주기 위한 목적으로 신의료기술평가에 대한 권고체계를 개선하고자 하였다. 건강보험 요양급여결정과정에서 사용되고 있는 가치와 과거 신의료기술로 인정된 기술에 대해 의료행위전문평가위원회에서 의사결정하였던 사례 분석을 토대로, 문헌

적 근거 이외 신의료기술평가 단계에서 보강해야 할 평가요인들을 추출하여 검토하였다.

신의료기술에 대한 급여 의사결정시에 중요하게 고려되었던 판단 근거로는 질병의 중중도, 치료효과성(임상적 유용성), 대체가능한 기술 유무, 사회적(정책적) 요구도, 편의성 증대, 남용 가능성, 장기추적관찰 필요, 비용효과성, 기 결정사항 반영이 고려된 것을 확인하였다. 또한, 유관기관 자문을 통해 신의료기술평가는 일차연구들의 체계적 문헌고찰을 통해 확인한 근거를 제시하였지만, 급여등재 결정시에는 해당 기술의 임상적 유용성 측면을 가장 우선순위로 보고 있으며, 의학교과서 및 가이드라인 등을 통해 검토하고 있는 사항을 확인하였다. 이에, 신의료기술평가 단계에서 임상적 유용성 측면에서의 근거자료로써 보완 가능한 항목으로, 중재가 환자의 의료결과에 미치는 영향력과, 의학교과서 및 임상가이드라인에서 확립되어 있는지, 실제 국외임상에서 사용되고 있는 현황으로 살펴보았다.

그 외, 신의료기술평가 단계에서 현재 평가시에도 고려는 하고 있으나 구체적으로 제시화하지 않았던 영역으로, 질병의 중증도, 대체기술가능 유무를 평가요인으로 추출하였다.

또한, 사례분석을 통해 고난이도 기술 또는 장기추적 관찰이 필요한 기술의 의결사항이나 향후 의료기술의 재평가와 연계를 고려하여, 신의료기술평가 당시 시점에서의 재평가 필요성에 대한 부분을 평가항목으로 포함하였다.

동 평가기준을 적용하여 권고등급을 결정하게 될 경우에는 급여결정시에 근거자료 보완이라는 측면에서 기존의 관련근거들이 의사결정을 내리기 애매한 경계선상상에 있을 때, 이를테면 선별급여 본인부담률 결정시 활용 가능하여 급여결정에 도움이 될 수 있다는 의견이 있긴 하였으나 전반적으로는 건강보험 의사결정시 검토되는 항목을 신의료기술평가단계에서 다룬다는 측면에서 기관 간의 업무가 중복되는 문제점이 있다는 비판이 제기되었다. 또한, 사례분석을 통해서도 확인되었듯이 급여결정은 문헌적 근거수준보다는 한정된 보건의료자원 내에서 비용효과성을 고려하여 판단하고 있기 때문에, 신의료기술평가시에 비용효과성을 고려하지 않은 채 다른 항목으로 추가적으로 검토한 권고 결과가 급여결정에 미치는 영향은 현재의 평가체계와 동일하게 미미할 것으로 보여 개선(안)에 대한 실효성은 부족하다는 의견이었다.

표 58. 신의료기술평가 권고등급 평가기준(1안)

<table>
<tr><th>검토항목</th><th>질문내용</th></tr>
<tr><td rowspan="3">임상적 유용성</td><td>• 중재(시술/검사)에 따른 환자의 의료결과에 미치는 영향력이 어떠한가?</td></tr>
<tr><td>• 의학교과서 및 가이드라인에서 동 기술의 임상적 유용성에 대해 언급하고 있거나 사용을 권고하고 있는가?</td></tr>
<tr><td>• 임상에서 사용되고 있는 현황(예: 국외 보험등재)은 어떠한가?</td></tr>
<tr><td>문헌적
근거수준</td><td>• 평가된 의료기술의 문헌적 근거수준은 어떠한가?
<table><tr><td>A</td><td>B</td><td>C</td><td>D</td><td>R (신속평가)</td></tr></table></td></tr>
<tr><td>기술의 확립성</td><td>• 중재시술의 경우, 시술방법에 있어 표준화되어 있는가?
• 진단검사의 경우, 검사방법에 있어 임계치 확립 및 결과가 신뢰할 수 있는가?</td></tr>
<tr><td>질병의 중증도/
희귀성/질병부담</td><td>• 질병 중증도/희귀성/질병부담이 높은 환자를 대상으로 하는 기술인가?</td></tr>
<tr><td>대체가능 기술 유무</td><td>• 대체 가능한 기술이 있는가?</td></tr>
<tr><td rowspan="2">재평가 필요성</td><td>• 실제 임상자료의 근거축적을 통해 의료기술의 재검증 필요성이 높은 기술인가?</td></tr>
<tr><td>• 기술의 난이도가 높아 국가 차원에서 관리가 필요한 기술인가?</td></tr>
</table>

나, 평가기준(2안)

신의료기술이 임상전문가들에 의해 적극 활용된다는 점을 고려하여 의과학적 근거에 대한 가치평가를 토대로 임상적 권고의 강도를 제시할 수 있도록 개선하고자 하였다. 이를 위해, 국외 주요 의료기술평가기관 및 임상진료지침을 개발하는 학회들에서 사용하고 있는 GRADE에서 제시하는 의사결정구조(Evidence to decision, EtD framework)를 적용하기 위한 방안을 검토하였다.

GRADE EtD는 문제의 우선순위(질병부담), 근거수준, 이득과 위해간의 비교(저울질), 비용, 환자의 가치와 선호, 자원사용(비용효과성), 형평성, 이해당사자들의 수용성, 실행가능성의 9가지 항목을 고려하여 보건의료 중재기술(행위)에 대한 임상적/공중보건 권고 및 급여결정 지원에 도움을 주기 위한 의사결정 구조이다.

본 연구에서는 신의료기술평가가 시장진입여부를 결정하는 단계에서의 평가라는 점을 감안하여, 신의료기술에 대한 권고등급을 결정하기 위해서 평가기술(중재) 관련 질병부담, 효과(긍정 효과, 부정 효과), 효과추정에 대한 신뢰도(근거수준), 가치, 효과간의 비교를 평가요소로 살펴보았다.

표 59. 신의료기술평가 권고등급 평가기준(2안)

항목	판단기준
문제	• 문제가 우선사항(priority)인가?
효과	• 예상되는 긍정 효과(desirable effects)는 얼마나 큰가?
	• 예상되는 부정 효과(undesirable effects)는 얼마나 작은가?
효과의 신뢰성 (근거수준)	• 효과에 대한 근거 신뢰성이 전반적으로 어떠한가?
가치	• 주요 의료결과에 대한 가치판단에 중요한 불확실성 또는 가변성이 있는가?
효과 간의 비교	• 긍정 효과와 부정 효과를 고려했을 때, 중재법과 비교법 중 어느쪽이 더 선호되는가?

문제의 우선순위는 평가기술(중재)과 관련한 질병의 심각성, 시급성, 질병 부담 등이 고려사항이 될 수 있다. 과거 신의료기술평가 사례에서도 임상적 근거 외에도 이러한 문제 우선순위를 고려하여 신의료기술로 인정되는 경우가 있었으며, 급여결정시에도 질병부담 등을 주요하게 보고 급여결정에 활용되는 것을 확인하였다.

GRADE는 체계적문헌고찰을 토대로, 결과에 대한 근거 요약 작성, 개별 근거의 질 결정, 결과의 중요도 결정, 전체적인 근거의 질 결정(근거수준), 효과 간의 비교(이득과 위해의 저울질) 등의 과정을 거쳐 권고의 강도를 결정한다.

기대하는 효과와 기대하지 않는 효과별로 구분하여 개별 결과에 대한 근거수준과 효과의 추정치를 확인하여 근거의 방향을 평가한다. 현재 신의료기술평가에서는 GRADEpro를 이용하여 근거수준을 결정하지는 않지만, 체계적문헌고찰을 통해 근거수준 판정에 필요한 자료를 충실히 다루고 있다고 생각되며, 향후 추가적으로 GRADEpro를 이용한 방법을 적용하여 근거 요약표를 생성하여 근거수준을 평가하는 것이 필요하겠다.

동 평가기준을 적용하는 것에 대해 신의료기술평가에 맞게 질문내용을 쉽게 바꾸고, 적용가능성을 확인하기 위해 좀 더 많은 시뮬레이션이 필요하다는 의견이 있었다. 또한, 권고강도를 제시할 때는 신의료기술로 인정되는 기술 뿐 아니라 연구단계기술에 있어서도 권고의 강도를 제시하는 것이 필요하다는 의견이 제시되었다.

표 60. 평가기준(안) 비교

구분	평가기준 제1안.	평가기준 제2안.
장점	· 급여결정시 고려되는 요소 일부를 사전에 항목별 검토하여 보험등재시 참고할 수 있는 근거자료 지원 가능	· 의사결정에 필요한 근거를 제공하는 범용된 방법론
단점	· 검토항목에 있어 기관 간 업무 중복성 · 평가항목이 권고등급(강도) 결정으로 도출되는 논리 근거가 빈약	· 평가방법이 보다 심층화되어, 평가기간 및 평가 인력 부족 · GRADE EtD 평가에 대한 교육 및 훈련 부족 · 근거수준(효과의 신뢰도) 판단을 위해서는 GRADE 평가방법 tool 고려 필요
	· 보험등재 연계에 대해 법적인 제도화가 이루어지지 않을 경우 권고등급 실효성 없음 (예 : 권고등급이 보험등재시 반영이 안 될 경우 문제제기 가능성) · 권고의 강도(높음, 낮음) 판단할 수 있는 명확한 기준 없이는 합의도출이 어려움	

4.2.3. 평가절차 및 합의방식

신의료기술평가 권고등급에 대한 의사결정을 하는 주체 및 합의방식에 대해 전문가 자문회의를 통해 논의하였다. 권고는 상기 평가요소를 고려해서 투명하고 일관성 있게 과학적으로 합의(appraisal)과정을 수행해야 하기 때문에, 권고를 제시하는 위원회는 객관적으로 의사결정 할 수 있는 위원으로 구성되어야 한다. 이런 점을 고려하여, 권고등급 결정은 현재의 신의료기술평가 절차(소위원회에서 검토의견을 제시하고, 분과위원회에서 확인 후 신의료기술평가위원회에서 최종 결정)로 수행하는 방안이 효율적일 것으로 검토하였다.

합의방식은 상기 제안한 권고등급 평가기준 개별 항목에 대해 논의 후, 숙의과정·합의를 통하여 권고등급을 결정하는 것으로 하였다.

다만, 신의료기술평가의 범위를 권고까지 포함하여 평가하기 위해서는 이를 위한

자료준비와 평가가 필요하다. 현재, 신의료기술평가의 법정기한이 250일로 단축되어 있는 상황에서 권고평가까지 포함하여 동시에 수행하기에는 시간적, 인력적 자원이 턱없이 부족한 상황이다. 이에 대한 대안으로, 신의료기술평가 결과 안전성 및 유효성이 있는 신의료기술로 심의되는 경우에 신의료기술평가 법정기한에 구애받지 않는 별도의 추가 절차를 마련하여 권고등급 평가를 진행하는 것이 방안이 될 수 있다. 또한, 안전성 유효성 평가 소위원회는 매번 위원의 구성이 달라지기 때문에 가치 등의 판단이 개입되는 권고등급 평가에 있어 일관성 있는 결과를 도출하는데 제한점이 있다. 권고평가를 진행하는 별도의 추가절차를 마련하게 될 시에는, 권고평가만을 위한 별도의 위원회 구성도 고려할 필요가 있다. 위원회 구성은 평가와 관련된 임상전문가(유동위원), 체계적문헌고찰 전문가, 통계전문가 등을 포함할 수 있으며, 신의료기술평가 권고 결과의 수용성 및 적용성 등을 고려할 때 관계부처 등의 정책결정자 또는 이해관계자 등을 포함할 수 있겠다. 하지만, 별도의 위원회(협의체)를 구성해서 진행할 시에는 신의료기술평가와 관련한 민원인 입장에서 업무처리 지연으로 민원제기 가능성이 있으며, 추가적인 시간을 투자한 만큼 동 권고등급의 활용성 및 수용성이 크지 않을 것으로 예상되어 신의료기술평가에 적용하기에는 어려움이 있을 것으로 보여지며, 동 권고등급의 활용성 등에 대한 관련부처 간의 협의가 선행되는 등 추가적인 검토가 필요하다.

5. '권고체계(안)' 시범 적용

본 연구에서 도출한 두가지 권고체계(안)에 대해 과거 신의료기술로 인정된 기술의 처치 및 시술과 진단 영역별로 시범 적용하여, 동 권고체계의 평가방법 등에 대한 적용가능성과 수정보완사항 등에 대해 확인하고자 하였다.

5.1. 항PLA2R IgG 항체검사 [효소면역분석법][간접면역형광분석법]

5.1.1. 기술개요

동 기술은 막성 신병증 (의심) 환자를 대상으로 혈청 또는 혈장에서 항 Phospholipase A2 receptor(PLA2R) IgG 항체를 효소면역분석법으로 정량측정하거나, 간접면역형광분석법으로 정성 또는 반정량으로 측정하여 원발성 막성 신병증을 진단하는데 도움을 주는 기술이다.

동 기술은 2017년 3월 22일 신청되어 2017년 제5차 신의료기술평가위원회(2017.5.26.)에서는 동 기술을 현재 건강보험요양급여비용에 등재되어있지 않는 안전성 및 유효성에 대한 평가가 필요한 신의료기술 평가대상으로 판단하여 신장내과 3인, 소아청소년과 2인, 진단검사의학과 2인 총 7인으로 구성된 소위원회에서 평가하였으며, 2017년 제10차 신의료기술평가위원회(2017.10.27.)에서는 소위원회 검토결과에 근거하여 최종심의한 결과 '안전성 및 유효성이 있는 기술'로 평가되었다. 이후 신의료기술의 안전성·유효성 평가결과에 대한 보건복지부 고시(제2017-221호) 및 급여(보건복지부 고시 제2018-241호, 2018.11.9.)로 결정되었다(누-811, 항PLA2R IgG 항체 [정밀면역검사]). 항PLA2R IgG 항체검사 [효소면역분석법][간접면역형광분석법]의 신의료기술평가 내용은 표 61과 같다.

표 61 . 신의료기술평가 내용(항PLA2R IgG 항체검사 [효소면역분석법][간접면역형광분석법])

<table>
<tr><th>기술명</th><th>항PLA2R IgG 항체검사 [효소면역분석법][간접면역형광분석법]</th></tr>
<tr><td>평가계획</td><td>① 검사목적: 막성 신병증 의심환자에서 원발성 막성 신병증 환자 진단(다른 질환과 감별)에 도움
• 대상환자: 막성 신병증 의심환자
• 중재검사: 항 PLA2R IgG 항체검사
• 참고표준검사: 조직병리검사+2차성 막성 신병증을 감별하기 위해 수행되는 검사
• 의료결과
· 진단정확성

② 검사목적 : 막성 신병증 환자에서 원발성 막성 신병증과 이차성 막성 신병증을 감별진단하는데 도움
• 사전검사 : 조직병리검사
• 대상환자 : 막성 신병증 환자
• 중재검사 : 항 PLA2R IgG 항체검사
• 참고표준검사 : 2차성 막성 신병증을 감별하기 위해 수행되는 검사
• 의료결과 : 진단정확성</td></tr>
<tr><td>선택문헌</td><td>총 27편
진단법평가연구 27편
질평가 2+ 7편, 2− 20편
원발성 막성 신병증과 비원발성 막성 신병증 환자 감별에 대해 보고한 연구 15편
원발성 막성 신병증과 이차성 막성 신병증 환자의 감별진단에 대해 보고한 연구 21편
검사방법: ELISA 9편, IFA 8편, 웨스턴 블럿 2편, ALBIA 1편, TRFIA 1편</td></tr>
<tr><td>유효성 결과</td><td>○ 원발성 막성 신병증과 비원발성 막성 신병증 환자 감별진단(15편)
<table>
<tr><th></th><th>Sn</th><th>Sp</th><th>PPV</th><th>NPV</th></tr>
<tr><td>전체(15편)</td><td>0.44~0.98</td><td>0.83~1.00</td><td>0.78~1.00</td><td>0.35~1.00</td></tr>
<tr><td>질평가 2+이상(3편)</td><td>0.63~0.88</td><td>0.86~1.00</td><td>0.78~1.00</td><td>0.61~0.93</td></tr>
<tr><td>ELISA(9편)</td><td>0.44~0.98</td><td>0.83~1.00</td><td>0.88~1.00</td><td>0.35~1.00</td></tr>
<tr><td>질평가 2+이상(2편)</td><td>0.65, 0.67</td><td>0.97, 1.00</td><td>0.96, 1.00</td><td>0.61, 0.72</td></tr>
<tr><td>IFA(8편)</td><td>0.52~0.81</td><td>0.90~1.00</td><td>0.87~1.00</td><td>0.48~0.84</td></tr>
</table>
○ 원발성 막성 신병증과 이차성 막성 신병증 환자 감별진단(21편)</td></tr>
</table>

<table>
<tr><td>기술명</td><td>항PLA2R IgG 항체검사 [효소면역분석법][간접면역형광분석법]
<table><tr><td></td><td>Sn</td><td>Sp</td><td>PPV</td><td>NPV</td></tr><tr><td>전체(21편)</td><td>0.25~1.00</td><td>0.64~1.00</td><td>0.81~1.00</td><td>0.17~1.00</td></tr><tr><td>질평가 2+이상(5편)</td><td>0.63~0.82</td><td>0.89~1.00</td><td>0.96~1.00</td><td>0.44~0.61</td></tr><tr><td>ELISA(10편)</td><td>0.25~0.98</td><td>0.64~1.00</td><td>0.88~1.00</td><td>0.17~0.87</td></tr><tr><td>질평가 2+이상(2편)</td><td>0.74,0.82</td><td>0.89, 0.94</td><td>0.96, 0.97</td><td>0.57, 0.61</td></tr><tr><td>IFA(8편)</td><td>0.48~1.00</td><td>0.64~1.00</td><td>0.81~1.00</td><td>0.25~1.00</td></tr><tr><td>질평가 2+이상(3편)</td><td>0.72~0.82</td><td>0.94~1.00</td><td>0.97~1.00</td><td>0.50~0.55</td></tr></table></td></tr>
<tr><td>소위원회
검토결과</td><td>ⅰ) 민감도 및 음성예측도의 수준은 다른 원인으로도 원발성 막성 신병증이 진단될 수 있는 점을 고려시 수용가능하며,
ⅱ) 특이도 및 양성예측도가 높게 보고되어 항 PLA2R IgG 항체가 질환 특이적인 측면에서 원발성 막성 신병증을 진단하는데 있어 임상적으로 도움이 될 수 있다는 의견</td></tr>
<tr><td>신의료기술평가
위원회 심의결과</td><td>안전성 및 유효성이 있는 기술로 심의 (근거의수준 C)</td></tr>
</table>

5.1.2. 평가항목 1안 적용시 권고등급 논의 결과

건강보험심사평가원에서 급여결정 시 기준으로 삼는 원칙들을 참고하여 만든 평가항목 1안을 적용하여 임상적 유용성, 문헌적 근거수준, 기술의 확립성, 재평가 필요성, 질병의 중증도/희귀성/질병부담, 대체가능기술 유무를 토대로 논의하여 종합의견을 제시한 결과는 다음과 같다.

ⅰ) 임상적 유용성과 관련해서는 동 검사가 원발성 막성 신병증 진단을 좀 더 수월하게 수행하는데 도움이 될 수 있으며, 교과서 및 가이드라인에는 아직 제시되어있지 않으나 최신 연구의 흐름상 임상적 유용성이 있다는 내용의 저널들이 발표되고 있고 미국에서도 사용되고 있어 임상적 유용성이 있다는 의견이었다.

ⅱ) 문헌적 근거수준의 경우 이미 평가한 결과에서 근거수준을 제시하고 있기 때문에 평가항목에서 포함시킬 필요가 없다는 의견 및 GRADE를 이용한 문헌적 근거수준을 제시하는 것이 필요하다는 의견이 있었다.

ⅲ) 기술의 확립성에 대해서는 검사분야에 있어서 기관마다, 사용되는 키트마다

방법 내용이 조금씩 다를 수 있기 때문에 확립되었다거나 표준화되었다는 말은 적절하지 않으며 '검사방법이 신뢰할 수 있는가?' 정도의 질문이 적절하다는 의견이었으며, 동 검사의 경우 신뢰할 수 있다는 의견이었다.

iv) 재평가 필요성과 관련하여 시장진입단계에서 동 질문항목이 적절하지 않다는 의견이 있었으며, 동 기술의 경우는 재평가 필요성은 없다는 의견이었다.

v) 질병의 중증도/희귀성/질병부담과 관련하여 동 검사의 경우 막성신병증 환자의 일부가 조기 치료가 이루어지지 않을 경우 중증신장질환으로 진행될 수 있기 때문에 중증도가 높은 질환이라는 의견이 있었으나 보는 관점에 따라 동 질문에 대한 의견은 다양할 수 있다는 의견이 있었다.

vi) 대체가능기술 유무와 관련하여서는 보는 관점에 따라 대체기술이 있는지 여부에 대해 의견이 다양할 수 있으며, 동 기술의 경우 현재 조직검사 및 2차성 막성 신병증을 배제하기 위해 수행되는 검사 외에 원발성 막성 신병증을 판단할 수 있는 혈액검사 수준에서는 다른 대체가능한 검사가 없다는 의견이 있었다.

종합적으로 동 기술의 경우 권고의 강도는 높다고 전원동의로 합의하에 의견이 제시되었으나, 위의 평가항목들을 종합해서 권고의 강도가 높고 낮음을 판단하게 되는 논리적 근거가 빈약하다는 의견이 있었다. 논의결과에 대한 상세내용은 표 62와 같다.

표 62. 권고등급 논의결과(항PLA2R IgG 항체검사, 평가항목 1안 적용)

검토항목	내용
임상적 유용성	1. 중재(시술/검사)에 따른 환자의 의료결과에 미치는 영향력이 어떠한가?
	기존에는 조직검사를 통해 진단을 내리고 있으나 간혹 악성종양이나 이차성 막성 신병증과의 감별진단이 필요하므로 다양한 검사를 통해 확인하나 동 검사가 있을 경우 좀더 진단을 수월하게 할 수 있음.
	2. 의학교과서 및 가이드라인에서 동 기술의 임상적 유용성에 대해 언급하고 있거나 사용을 권고하고 있는가?
	2012년 이후로 가이드라인이 업데이트 되지 않아 동 검사에 대한 권고내용이 포함되어있지 않으나 최근에 나온 문헌들에서 동 검사의 임상적 유용성을 입증하는 내용들이 많이 나오고 있는 추세임. 의학교과서나 가이드라인에 제시안되어있다고 임상적 유용성이 없다고 말할 수는 없으며, 최신 연구 흐름에 대해 전문가 자문을 제시할 수 있을 것이라는 의견제시

검토항목	내용
	3. 임상에서 사용되고 있는 현황(예: 국외 보험등재)은 어떠한가?
	사용되고 있음
문헌적 근거수준	3. 평가된 의료기술의 문헌적 근거수준은 어떠한가?
	문헌적 근거수준을 판단항목에서 배제해도 되겠다는 의견 및 GRADE 도구를 활용하는 것도 필요하다는 의견 제시
기술의 확립성	4. 검사방법에 있어 표준화가 확립되어 있는가?
	표준화라는 말은 적합하지 않으며 '검사방법이 신뢰할 수 있는가?' 로 질문을 바꾸는 것이 낫겠다는 의견이 있었으며 동 검사의 경우 검사방법을 신뢰할 수 있다고 판단함.
재평가 필요성	5. 실제 임상자료의 근거축적을 통해 의료기술의 재검증 필요성이 높은 기술인가?
	질문항목을 ' 실제 임상자료의 근거의 축적을 통해 의료기술의 임상적 유용성에 대한 재검증 필요성이 높은 기술인가' 로, 임상적 유용성에 대한 재검증이라는 문장을 넣는 것이 낫겠다는 의견 및 시장진입단계에서 인정된 기술에 대해 동 질문항목을 넣을 필요는 없다는 의견이 있었음. 동 기술의 경우 재검증 필요성은 없다는 의견이었음.
	6. 기술의 난이도가 높아 국가 차원에서 관리가 필요한 기술인가?
	신의료기술로 인정된 기술에 대해 국가 차원에서 관리가 필요한 기술인지 질문하는 것이 적절한지에 대한 의견이 있었으며, 동 기술의 경우 국가차원의 관리는 필요없다는 의견이 있었음.
질병의 중증도/희귀성/질병부담	7. 질병 중증도/희귀성/질병부담이 높은 환자를 대상으로 하는 기술인지?
	동 검사대상의 경우 질병부담이 높고 중증하다고 판단함
대체가능 기술 유무	8. 대체 가능한 기술이 있는가?
	검사의 특성을 봤을 때 대체가능한 기술은 없다는 의견 및 동 검사의 경우 '해당사항 없음' 이 적절하다는 의견 제시
종합의견	권고강도: 높음(강함)/낮음(약함) : 높음 – 권고강도를 합의를 통해서 높음으로 도출은 할 수 있으나 평가항목 각 질문들의 답을 통해서 권고강도를 높음으로 논리적으로 도출하는 근거가 빈약하다는 의견이 있었음.

5.1.3. 평가항목 2안 적용시 권고등급 논의 결과

GRADE EtD를 활용한 평가항목을 적용하여 문제, 검사정확성, 검사정확성에 대한 근거의 확실성, 검사의 효과에 대한 근거의 확실성, 검사결과/관리에 대한 근거의 확실성, 효과에 대한 확실성, 가치, 효과에 대한 균형을 토대로 논의하여 종합의견을 제시한 결과는 다음과 같다.

ⅰ) 문제와 관련해서는 평가항목 1안에서의 질병의 중증도/희귀성/질병부담 항목과 동일한 사유로 문제의 우선순위에 포함된다는 의견이었다.

ii) 검사정확성, 검사정확성에 대한 근거의 확실성, 검사의 효과에 대한 근거의 확실성, 검사결과/관리에 대한 근거의 확실성, 효과에 대한 확실성과 관련해서는 동 기술의 경우 민감도가 낮고 특이도가 높은 특성을 가지나 질환 특이적인 측면에서 동 검사가 임상적으로 유용하다는 내용을 진단정확성의 측면에서 높다, 낮다라고 판단하기 어렵다는 의견이 있었다. 그 외에 판단기준에 보통이 포함되어 있어, 대부분 보통을 체크하게 될 가능성이 높으며, 그럴 경우 최종적으로 권고등급을 제시할 때 도출되는 논리적인 근거가 부족하다는 의견을 제시하였다. 종합적으로 해당 항목에 대해서는 판단하기 어렵다는 의견(전원동의)를 제시하였다.

iii) 가치와 관련해서는 동 기술의 경우 변동성 및 불확실성은 낮다고 판단하였다.

iv) 효과에 대한 균형과 관련해서는 비교검사와 비교된 연구결과가 없기 때문에 판단하기 어렵다는 의견을 제시하였다.

v) 최종 권고등급은 높음으로 전원 동의로 합의가 되었으나 상기 제시된 평가항목에서 권고등급 높음을 도출하는 논리적 근거를 설명하기 어려워 동 평가항목을 활용하는데 어려움이 있다는 의견이 있었다. 상세 내용은 표 63과 같다.

표 63. 권고등급 논의결과(항PLA2R IgG 항체검사, 평가항목 1안 적용)

항목	의견	결론
문제	우선순위가 높다고 판단함.	예
검사정확성 검사정확성에 대한 근거의 확실성	· 동 기술의 경우 민감도가 낮고 특이도가 낮은 상황에서 GRADE의 판단기준을 체크하기가 어렵다는 의견 제시 · 판단기준에 보통이 포함되어있기 때문에 대부분 보통으로 체크를 하게 될 경우 권고등급의 높고 낮음을 도출할 수 있는 근거의 논리가 희박할 수 잇다는 의견 제시	판단하기 어려움

항목	의견	결론
검사의 효과에 대한 근거의 확실성 검사결과/관리에 대한 근거의 확실성 효과에 대한 확실성	· 매우 낮음, 매우 부정확함 및 부정확함, 낮음은 아니나, 보통, 높음 혹은 정확함, 매우 정확함을 판단할 수 있는 근거가 부족하다는 의견	
가치	변동성 및 불확실성은 낮다고 판단함	변동성 및 불확실성 낮음
효과에 대한 균형	비교검사가 없기 때문에 판단하기 어렵다는 의견 제시	모르겠음
최종 권고등급	동 기술의 경우 권고의 강도를 높음으로 제시할 수 있으나 위의 근거에서 높음으로 도출할 수 있는 논리적 근거를 설명하기 어렵다는 의견 제시	높음

5.2. 툴륨레이저를 이용한 전립선 광적출술

5.2.1. 기술개요

툴륨레이저를 이용한 전립선 광적출술은 양성 전립선 비대증 환자를 대상으로 툴륨레이저를 이용하여 비대해진 전립선에 대해 피막을 따라 절제 후 조직분쇄기를 이용하여 적출함으로서 전립선 비대증을 치료하는 수술방법이다.

동 기술은 2013년 9월 6일 신청되어 2013년 제10차 신의료기술평가위원회(2013.11.1.)에서는 동 기술을 현재 건강보험요양급여비용에 등재되어있지 않는 안전성 및 유효성에 대한 평가가 필요한 신의료기술 평가대상으로 판단하여 비뇨기과 3인, 가정의학과 1인 총 4인으로 구성된 소위원회에서 평가하였으며, 2014년 제3차 신의료기술평가위원회(2014.3.28.)에서는 소위원회 검토결과에 근거하여 최종심의한 결과 '안전성 및 유효성이 있는 기술'로 평가되었다. 이후 신의료기술의 안전성·유효성 평가결과에 대한 보건복지부 고시 제2014-68호(2014.5.14.)로 공표되었다. 툴륨레이저를 이용한 전립선 광적출술의 신의료기술평가내용은 표 64와 같다.

표 64. 신의료기술평가 내용(툴륨레이저를 이용한 전립선 광적출술)

<table>
<tr><th>기술명</th><th>툴륨레이저를 이용한 전립성 광적출술</th></tr>
<tr><td>평가계획</td><td>• 대상환자: 양성 전립선 비대증 환자
• 중재수술: 툴륨레이저를 이용한 전립선 광적출술
• 비교수술:
· 경요도적 전립선절제술(Trans-Urethral Resection of the Prostate, TURP),
· 개복전립선적출술(Open Prostatectomy, OP)
· 레이저전립선수술(광응고술, 광적출술, 기화술, 절제술)
– Holmium: YAG(Ho: YAG)
– Potassium Titanyl Phosphate(KTP)
• 의료결과
· 안전성
수술 관련 합병증 및 부작용
· 유효성
요역동학적 지표개선정도 : 최대요속, 배뇨 후 잔뇨량
증상호전정도 : 국제전립선증상점수
삶의 질
수술 관련 요인 : 도뇨관 삽입기간, 수술시간, 재원기간, 조직절제율</td></tr>
<tr><td>선택문헌</td><td>총 12편
무작위임상시험연구 2편(TURP와 비교 1편, HoLEP와 비교 1편)
전후 연구 10편</td></tr>
<tr><td>안전성 결과</td><td>2편의 무작위 임상시험연구(각각 경요도적 전립선 절제술, 홀뮴레이저를 이용한 전립선 광적출술과 비교)에서 중재시술이 비교시술에 비해 출혈 관련 지표(혈액손실, 헤모글로빈감소)가 유의하게 낮았으며, 그외 합병증에서는 차이가 없었음

· TURP와 비교
<table><tr><th colspan="2">합병증종류</th><th>TuLEP(79명)</th><th>TURP(79명)</th><th>p</th></tr><tr><td colspan="2">TURS</td><td>없음</td><td>없음</td><td>–</td></tr><tr><td colspan="2">요정체</td><td>1명</td><td>3명</td><td>–</td></tr><tr><td colspan="2">요도, 방광경부 협착</td><td>없음</td><td>없음</td><td>–</td></tr><tr><td rowspan="2">출혈</td><td>헤모글로빈 감소정도</td><td>0.15±0.02</td><td>0.30±0.03</td><td>0.045</td></tr><tr><td>수혈</td><td>없음</td><td>없음</td><td>–</td></tr></table>
· HoLEP와 비교
<table><tr><th colspan="2">합병증 종류</th><th>TuLEP(71명)</th><th>HoLEP(62명)</th><th>p</th></tr><tr><td colspan="2">피막천공</td><td>1명</td><td>1명</td><td>–</td></tr><tr><td rowspan="2">출혈</td><td>헤모글로빈 감소정도(g/dl)</td><td>0.5±0.06</td><td>0.5±0.10</td><td>0.107</td></tr><tr><td>혈액손실</td><td>130.1±20.3</td><td>166.6±17.1</td><td>0.045</td></tr></table></td></tr>
</table>

기술명	툴륨레이저를 이용한 전립성 광적출술

합병증 종류	TuLEP(71명)	HoLEP(62명)	p
수혈	없음	없음	–
전해질 불균형으로 인한 질환 발생	없음	없음	–
혈청 Na 감소정도(mmol/l)	7.6±2.0	6.6±2.7	0.082
경미한 실금	없음	1명	–

· 전후 연구 10편

지표(연구편수)	결과
수혈(10편)	1~2.7%
재수술률(7편)	2.2~7.2%
지속적인 스트레스성 요실금 및 지속적 과민성 요증후군(5편)	1~4.76%

유효성 결과

○ 비교군 연구:

– 요역동학적 효과, 증상호전 정도, 삶의 질에서 비교시술과 모두 유사한 수준

– 수술시간은 툴륨레이저를 이용한 전립선 광적출술이 비교군에 비해 유의하게 길었으나 도뇨관삽입기간, 입원기간에서는 경요도적 전립선 절제술에 비해 유의하게 짧았으며 홀뮴레이저를 이용한 전립선 광적출술과는 유사한 것으로 보고

○ TURP와 비교

유효성지표			ThuLEP	TURP	p
요역동학적 지표 개선정도	Qmax (ml/s)	수술전	8.7±2.8	9.1±3.2	.143
		18개월후	22.9±12.7	23.2±13.3	.289
	PVR (ml)	수술전	79.5±29.3	72.4±28.1	.123
		18개월후	30.7±15.2	31.2±14.3	.182
증상호전정도(IPSS)		수술전	22.7±4.3	23.4±3.7	.081
		18개월 후	5.7±2.1	5.2±1.9	.131
QOLS		수술전	3.9±1.2	4.9±1.3	.127
		18개월 후	1.2±1.1	1.2±0.9	.119
수술관련 요인	수술시간(분)		65.4±22.2	47.4±15.9	.022
	도뇨관 삽입기간(일)		2.1±0.8	3.5±1.2	.031
	방광세척양(L)		12.4±6.4	27.2±5.2	.045
	입원기간(일)		2.5±1.4	4.6±1.4	.026

ThuLEP, Thulium Laser Enucleation of the Prostate; HoLEP, holmium laser trasurethral enucleation of the prostate; Qmax, maximum flow rate; PVR, postvoid residual urine volume; IPSS, International Prostate Symptom Score; QOLS, Quality of Life Score(IPSS의 8번째 질문으로 0(delighted)~6(terrible)로 점수화됨)

<table>
<tr><th>기술명</th><th>툴륨레이저를 이용한 전립성 광적출술</th></tr>
<tr><td></td><td>
○ HoLEP와 비교
<table>
<tr><th colspan="3">유효성 지표</th><th>ThuLEP</th><th>HoLEP</th><th>p</th></tr>
<tr><td rowspan="4">요역동학적 지표 개선 정도</td><td rowspan="2">Qmax</td><td>개선정도</td><td>향상</td><td>향상</td><td rowspan="2">언급없음</td></tr>
<tr><td>p</td><td>유의</td><td>유의</td></tr>
<tr><td rowspan="2">PVR(ml)</td><td>개선정도</td><td>11.3</td><td>11.8</td><td rowspan="2">언급없음</td></tr>
<tr><td>p</td><td>유의</td><td>유의</td></tr>
<tr><td colspan="2" rowspan="2">IPSS</td><td>개선정도</td><td>5.2</td><td>6.2</td><td rowspan="2">언급없음</td></tr>
<tr><td>p</td><td>유의</td><td>유의</td></tr>
<tr><td colspan="2" rowspan="2">QOLS</td><td>개선정도</td><td>향상</td><td>향상</td><td rowspan="2">언급없음</td></tr>
<tr><td>P</td><td>유의</td><td>유의</td></tr>
<tr><td rowspan="3">수술관련 요인</td><td colspan="2">수술시간(분)</td><td>72.4±19.4</td><td>61.5±20.2</td><td>0.034</td></tr>
<tr><td colspan="2">도뇨관 삽입기간(일)</td><td>2.4±1.0</td><td>2.5±1.0</td><td>0.118</td></tr>
<tr><td colspan="2">절제된 조직무게(g)</td><td>37.6±12.1</td><td>40.4±10.6</td><td>0.072</td></tr>
</table>
ThuLEP, Thulium Laser Enucleation of the Prostate; HoLEP, holmium laser trasurethral enucleation of the prostate; Qmax, maximum flow rate; PVR, postvoid residual urine volume; IPSS, International Prostate Symptom Score; QOLS, Quality of Life Score

○ 단일군 연구 10편: 요역동학적 효과, 증상호전정도, 삶의 질 모두 수술 전후 유의하게 개선된 것으로 나타남
</td></tr>
<tr><td>소위원회 검토결과</td><td>
○ 안전성: 동 기술의 합병증이 경요도적 전립선 절제술 및 홀뮴레이저를 이용한 전립선 광적출술과 유사한 수준으로 수용가능하다고 평가
○ 유효성: 요역동학적 효과, 증상호전정도, 삶의 질에서 경요도적 전립선 절제술 및 홀뮴레이저를 이용한 전립선 광적출술과 비교시 유사한 수준의 임상적 치료효과를 보여 유효성이 있다고 판단
</td></tr>
<tr><td>신의료기술 평가위원회 심의결과</td><td>안전성 및 유효성이 있는 기술로 심의(권고등급 B)</td></tr>
</table>

5.2.2. 평가항목 1안 적용시 권고등급 논의 결과

건강보험심사평가원에서 급여결정 시 기준으로 삼는 원칙들을 참고하여 만든 평가항목 1안을 적용하여 임상적 유용성, 문헌적 근거수준, 기술의 확립성, 재평가 필요성, 질병의 중증도/희귀성/질병부담, 대체가능기술 유무를 토대로 논의하여 종

합의견을 제시한 결과는 다음과 같다.

ⅰ) 임상적 유용성과 관련해서는 동 시술이 증상개선에 효과가 있기 때문에 환자에게 긍정적인 영향을 주고 있으며, 가이드라인 및 국외에서 사용되고 있는 현황을 확인하였을 때 임상적 유용성이 있다는 의견이었다.

ii) 문헌적 근거수준의 경우 이미 평가한 결과에서 근거의 수준을 B로 제시하고 있어서 해당 내용을 제시하였다.

iii) 기술의 확립성과 관련해서는 동 평가항목이 권고의 강도를 제시하기 위해서 시술방법의 표준화라는 용어가 적절하며 필요하다는 의견 및 시장진입단계에 있는 신의료기술평가에 있어서 표준화라는 용어가 부담스럽다는 의견으로 전문가간 이견이 있었으며, 동 기술의 경우 전원 동의로 시술방법에 있어서 표준화가 된 기술로 볼 수 있다는 의견이었다.

iv) 재평가 필요성과 관련하여 시장진입단계에서 동 질문항목이 적절하지 않다는 의견이 있었으며, 동 기술의 경우는 문헌적 근거가 무작위 임상시험 연구 2편으로 근거의 양이 충분하지 않아 재검증이 필요하다는 의견이었다.

ⅴ) 질병의 중증도/희귀성/질병부담과 관련하여 동 기술의 경우 질병부담이 높은 환자로 해당 질문항목이 권고의 강도를 결정하는데 있어 필요한 질문이라는 의견 및 해당질문은 나중에 급여결정시 필요한 항목으로 권고 강도를 결정할 때는 불필요한 질문이라는 의견으로 해당질문의 필요성에 대해 전문가간 이견이 있었다.

vi) 동 기술은 전원 동의로 대체가능한 기술이 있는 기술로 볼 수 있다는 의견으로 권고의 강도를 결정할 때 동 질문 항목의 필요성에 대해 필요하다는 의견 및 동 질문은 나중에 급여결정할 때 필요하고 권고 강도를 결정할 때는 필요없다는 의견으로 전문가간 이견이 있었다.

종합적으로 동 기술의 경우 권고의 강도는 중등도에 가까운 낮음이라는 의견과 중등도에 가까운 높다라는 의견으로 전문가간 이견이 있었으며, 건강보험심사평가원에서 급여결정시 확인하는 사항과 중복된다는 의견이 일부 있었다. 논의결과에 대한 상세내용은 표 65와 같다.

표 65. 권고등급 논의결과(툴륨레이저를 이용한 전립선광적출술, 평가항목 1안 적용)

검토항목	내용
임상적 유용성	1. 중재(시술/검사)에 따른 환자의 의료결과에 미치는 영향력이 어떠한가?
	증상개선에 도움이 안되는 기술을 신의료기술로 인정해줄 수 없기 때문에 동 질문항목에는 전원동의로 환자에게 긍정적 영향력을 준다는 의견 제시
	2. 의학교과서 및 가이드라인에서 동 기술의 임상적 유용성에 대해 언급하고 있거나 사용을 권고하고 있는가?
	전원동의로 가이드라인에서 동 기술을 고려할 수 있는 것으로 제시하고 있다는 의견 제시
	3. 임상에서 사용되고 있는 현황(예: 국외 보험등재)은 어떠한가?
	전원동의로 해외에서 사용되고 있다고 제시함 - 추가적으로 해외에서는 툴륨레이저보다는 홀뮴레이저 사용에 대해 언급을 더 많이 하고 있으나 동 기술은 홀뮴레이저든 툴륨레이저든 전립선 광적출술이라는 기술에 초점을 두고 보는 것이 맞아 별도의 다른 기술로 볼 필요는 없다는 의견이 있었음. - 레이저에 따라 다른 기술로 볼 수 있을지 여부는 전문가간 이견이 있을 가능성은 있다는 의견이었음. - 레이저 종류는 의료진들이 임상에서 사용시 선택할 수 있는 영역으로 둘 수 있다는 의견이었음.
문헌적 근거수준	3. 평가된 의료기술의 문헌적 근거수준은 어떠한가?
	전원동의로 근거의 수준 B 제시
기술의 확립성	4. 시술방법에 있어 표준화가 확립되어 있는가?
	- 수술 테크닉에 있어 표준화는 분명히 있으며, 동 기술의 확립성을 확인하기 위해서 표준화라는 용어를 사용하는 것이 필요하다는 의견 및 시장진입 단계에 있는 신의료기술평가에 있어서 표준화라는 용어가 부담스럽다는 의견이 있었음. 동 기술의 경우는 표준화된 방법을 사용하는 기술인 부분에는 전원동의함. - 기술의 확립성 외에 윤리적인 측면을 질문하기 위해 ‘ 동 기술이 의료윤리적으로 문제가 없는 기술인가’ 라는 질문에 대한 필요성이 제시되었으며, 동 질문은 신의료기술 평가 당시 필요한 질문이나, 만일 권고등급에 대한 질문을 최종심의결과 이전에 소위원회 검토당시 하게 될 경우 의료윤리에 대한 질문은 할 수 있다는 의견이 있었음.
재평가 필요성	5. 실제 임상자료의 근거축적을 통해 의료기술의 재검증 필요성이 높은 기술인가?
	툴륨레이저를 이용한 전립선 광적출술의 경우 무작위임상시험연구가 2편밖에

검토항목	내용
	안되서 근거의 양이 충분하지 않기 때문에 재검증 필요성이 있는 기술이라는 의견이었음(전원동의).
	6. 기술의 난이도가 높아 국가 차원에서 관리가 필요한 기술인가?
	전원 동의로 동 질문은 권고등급을 결정할 때 동 질문은 필요없다는 의견이 제시되었음.
질병의 중증도/희귀성/질병부담	7. 질병 중증도/희귀성/질병부담이 높은 환자를 대상으로 하는 기술인지?
	동 기술은 질병부담이 높은 질환으로 동 질문항목은 적절하다는 의견 및 동 기술이 질병부담이 높은 기술이기는 하지만 권고의 강도를 결정할 때 동 질문은 필요하지 않으며 해당질문은 나중에 급여결정할 때 필요하다는 의견으로 동 질문항목의 필요성에 대해서는 전문가간 이견이 있었음
대체가능 기술 유무	8. 대체 가능한 기술이 있는가?
	전원 동의로 대체가능한 기술이 있는 기술이며, 동 질문이 권고 강도를 결정할 때 필요하다는 의견 및 동 질문은 나중에 급여결정할 때에 필요하고 권고강도를 결정할 때는 필요없다는 의견으로 이견이 있었음
종합의견	권고강도: 합의되지 않음 1인: 낮음, 대체가능한 기술이 있으며 임상데이터가 많이 축적되어있지 않음. 중에 가까운 낮음으로 볼 수 있음. 1인: 높음, 레이저 종류가 중요한 게 아니고 테크닉이 중요하므로 높다고 판단하였으며, 중에 가까운 높음으로 판단함.

5.2.3. 평가항목 2안 적용시 권고등급 논의 결과

GRADE EtD를 활용한 평가항목을 적용하여 문제, 효과, 효과의 확신도(근거수준), 가치, 효과간 비교를 토대로 논의하여 종합의견을 제시한 결과는 다음과 같다.

평가항목 2안을 적용 시에는 각자 평가항목별 체크박스에 대해 체크한 결과를 공유하는 방식으로 이루어졌으며, 해당 결과를 체크한 사유에 대해 서로 논의가 이루어지지는 않았다. 항목별로 '기대하지 않은 효과', '주요 의료결과에 대한 가치 판단에 불확실성이나 가변성이 존재하는가', '기대하는 효과와 기대하지 않는 효과를 고려 시 중재법이 비교법보다 더 선호되는가'에 대해 의견일치가 이루어졌으며, '문제가 우선순위에 해당되는가?', '기대하는 효과는 어느 정도인가',

‘효과에 대한 근거의 확실성은 전반적으로 어떠한가’ , 에 대해서는 전문가간 이견이 있었으나 크게 차이가 나지는 않았다. 최종 권고등급은 평가항목 1안을 적용했을 때와 동일한 결과를 제시하였으며 전문가간 동 평가항목이 더 편하게 체크할 수 있어서 쉽다는 의견과 평가항목 1안이 좀 더 구체적이고 이해하기 쉽다는 의견이 있었다. 논의결과에 대한 상세내용은 표 66과 같다.

표 66. 권고등급 논의결과(툴륨레이저를 이용한 전립선광적출술, 평가항목 2안 적용)

질문내용	결론
문제가 우선순위에 해당되는가?	전문가 간 이견 있음(아닐것 같음/그럴 것 같음)
기대하는 효과는 어느 정도인가?	전문가 간 이견 있음(보통/큼)
기대하지 않은 효과는 어느 정도인가?	전원동의: 작음
효과에 대한 근거의 확실성은 전반적으로 어떠한가?	전문가 간 이견 있음(보통/낮음)
주요 의료결과에 대한 가치판단에 불확실성이나 가변성이 존재하는가?	전원동의: 아마도 중요한 불확실성 또는 가변성이 없을 것임
기대하는 효과와 기대하지 않는 효과를 고려시, 중재법이 비교법보다 더 선호되는가?	전원동의 : 아마도 비교법이 더 선호
최종 권고등급	전문가 간 이견있음(낮음, 중에 가까운 낮음으로 볼 수 있음/높음, 중에 가까운 높음으로 판단함).

Ⅳ 결론 및 제언

신의료기술평가는 체계적 문헌고찰을 기반으로 하여, 사전에 정해진 평가계획에 따른 문헌 검색·선택, 선정된 문헌에 대한 비뚤림 위험 평가, 자료추출 등의 엄격하고 객관적인 연구 과정을 통해 도출된 문헌적 근거를 평가하여 건강보험에 등재가 되지 않은 (신)의료기술의 등재 목적을 위한 형태로 임상도입에 관한 의사결정을 수행하고 있다. 신의료기술의 평가결과는 권고의 방향이 신의료기술 승인과 연구단계기술로 구성되어 있으며, 동시에 체계적문헌고찰이 수행되는 경우에 근거수준이 함께 제시하는 형태로 결정되어지고 있다. 신의료기술로 승인된 기술의 평가결과는 건강보험과정안에서의 급여기준 마련을 위해 의학적 타당성, 치료효과성 등을 평가하는데 기초적인 문헌적 근거자료로써 검토되고 있다.

현재, 신의료기술평가에서 체계적 문헌고찰시 사용되고 있는 SIGN의 평가 도구는 연구의 설계에 따른 잠재적 비뚤림에 대한 가능성을 평가하고, 그에 따른 근거수준(강도)(원문[50])에서는 Grade of Recommendation으로 제시)을 A~D까지 등급화하고 있다. 그동안 신의료기술평가를 수행하면서 동 등급은 임상적 권고의 강도로 오해해 등급이 A로 높을수록 보험 등재시 급여 결정에 긍정적인 영향을 미칠 것으로 잘못 이해되기도 하였다. 하지만, 이는 연구설계에 따른 연구의 질을 기준으로 일관성, 적용가능성을 고려하여 결정된 등급으로써 근거에 대한 효과의 크기를 평가하지 않는 등 연구설계에 따른 근거수준이 직접적으로 반영된 등급이라는 제한점이 있었다. 또한, 의료기술평가와 관련한 방법론이 진보하면서 연구주제마다 연구설계 및 해석법이 다양하고 정교해짐에 따라 연구설계가 반드시 임상적 중요성을 반영하지 않는다는 점이 강조되면서 현재는 SIGN에서도 2013년부터 권고등급에 대한 판단을 GRADE 체계로 받아들여 사용하고 있다.

실제로, 신의료기술평가제도 도입 이후 평가가 수행된 총 1,279건 중 신의료기술평가

50) source : SIGN 50: A guideline Developer Handbook, 2011

사례 분석을 통해서 그 간의 신의료기술로 인정된 기술(총 800건)의 근거수준별로 건강보험 의사결정에 미치는 영향을 분석했을 때, 근거수준 평가가 이루어진 345건과 미수행 266건의 총 611건 중 급여(급여·선별급여) 결정(총 403건, 급여 366건·선별급여 37건) 비율은 근거수준 A 2%, B 10%, C 26%, D 17%, 미평가 45%이었고, 비급여 결정(총 208건) 비율은 근거수준 A 1%, B 9%, C 25%, D 24%, 미평가 41%으로, 근거수준에 따라 급여와 비급여 등재 비율이 유사하여 급여 등재가 신의료기술평가의 근거수준 등급과는 별개로 결정되는 것을 확인할 수 있었다.

또한, 신의료기술평가 결과를 기술유형별로 구분하여 근거수준 범주를 비교하였을 때, 처치 및 시술은 근거수준 A 6%, B 25%, C 13%, D 48%, 미수행 8%이었으며, 체외진단검사는 근거수준 A 0.3%, B 1%, C 22%, D 13%, 미수행 63%, 영상검사는 근거수준 A 0%, B 7%, C 56%, D 33%, 미수행 4%, 기타검사 근거수준 A 0%, B 3%, C 50%, D 37%, 미수행 10%로, 중재시술의 영역에서 근거수준 A, B, D의 비율이 비교적 높고, 반면 체외진단검사 등 검사 분야에서는 신속검토로 인해 미수행과 C, D의 비율이 높은 순으로 나타났다. 이는 체외진단검사 분야에 있어 그동안 제도개선의 영향으로 신속검토에 의한 비중이 높아진 것과, 연구설계에 따라 근거수준이 직접적으로 반영되었음을 다시 확인할 수 있다. 여기에서, SIGN의 근거수준 체계를 중재시술 분야와 진단 연구를 분리해서 맞출 필요가 있다. 그러기 위해서는 SIGN의 체계를 modified하여 사용하거나, 다른 방법론 지침(CEBM 체계, GRADE 접근법 등)을 고려해 볼 수 있으며 이를 적용하기 위해서는 후속 연구를 통해서 검토할 필요가 있겠다.

본 연구에서는 SIGN의 근거수준 체계를 유지하면서, 신의료기술평가 결과의 급여결정과 의료기술평가 전주기적 관점에서 재평가 대상 선별을 위한 합리적인 의사결정하는데 도움을 주기 위해 신의료기술 권고강도에 대한 등급화 하는 방안을 제시하고자 하였다. 그동안 신의료기술로 승인되어 시장진입된 이후 급여등재시 의사결정되었던 사례 분석과 의료기술평가기관에서 사용되는 체계 등을 참고하여 권고등급을 결정하기 위한 2개의 평가기준(안)을 제시하고, 이에 따른 권고등급 체계를 제안하였다.

신의료기술평가 결과, 신의료기술로 승인된 경우에 있어 권고의 강도를 강함과 약함의 2단계로 등급화하고, 등급을 결정함에 있어 신의료기술에 대한 급여 의사결정시에 중요하게 고려되었던 판단근거요소를 위주로 한 평가기준(1안)과 GRADE의 근거기반 의사결정 틀(Evidence to Decision framework)을 활용한 평가기준(2안)으로 접근하였다.

권고등급과 평가기준(안) 마련을 위해 임상전문가 및 이해관계자와의 논의 시, 신의료기술평가의 권고등급이 어떻게 활용될 것인지 그 목적성에 대한 의견들이 제시되었다. 신의료기술 권고등급이 급여결정에의 연계 활성화를 도모하기 위한 연구의 취지 및 목적에 대해서, 안전성 및 유효성의 평가 근거자료와 실제 정책적 결정과정 사이에는 실제 기관간의 평가관점 차이가 존재하기 때문에 신의료기술평가의 범위가 확대되어 여러 가치 요인들을 고려한 권고등급안을 제시하더라도 급여결정 시 고려되는 재정 및 보험정책적 요인들이 고려되지 않는 이상 급여 의사결정에 미치는 영향은 미미할 것으로 보인다는 것이 전반적인 의견이었다.

이에, 권고등급의 활용목적 및 평가기준(안)에 대해, 급여결정에 영향을 주기 위함보다는 임상에서 활용 가치 측면에서 권고강도를 접근할 필요가 있으며, 의사결정 내용에 있어서도 기관 간의 업무의 차별화와 의료기술평가에 보다 적합한 관점에서 평가기준 2안이 보다 합리적이라는 의견이었다. 다만, 평가기준 2안을 적용하기 위해서도 EtD 틀을 좀 더 깊이있게 사용할 필요가 있으며 신의료기술평가에 맞게 질문내용을 수정보완하고, 적용가능성에 대해서도 보다 많은 사례에 대해 시범적용 등의 후속연구가 필요하다는 의견이었다.

본 연구에서는 신의료기술평가 보고서의 평가결과 등급을 근거수준 이외의 다른 요인들을 고려한 권고등급안을 제시하였으나, 이를 신의료기술평가 과정에서 적용하기 위해서는 아직 더 검토해야 할 요인들이 많이 있다. 신의료기술평가에 보다 적합한 평가기준에 대해 전문가 및 이해관계자의 의견수렴을 통해 결정해야 하고, 결정된 평가기준(항목)에 대한 많은 시범적용을 통해 타당도 검증이 이루어져야 사용될 수 있을 것이다. 또한, 본 연구진에서는 신의료기술로 인정된 경우에만 권고강도를 제시하는 것으로 마련하였으나, 실제 안전성 유효성 평가 소위원회를 운영하면서 권고의 강도에 대한 평가가 이루질 것을 감안하면 신의료기술로 인정되는 기술 뿐 아니라 연구단계기술에 있어서도 권고의 강도가 제시될 필요성에 대해서도 추후 검토가 이루어져야 할 것이다.

동 평가내용은 기존의 신의료기술 문헌적 근거 평가 이외 다양한 다른 가치 항목들을 포함하여 평가가 필요하므로, 신의료기술평가 법정기한이 단축되어진 상황에서 평가를 위한 시간적, 인력적 자원이 부족하여 이를 적용하기 위해서는 제도적 보완이 필요하다. 그 대안으로는 신의료기술로 인정된 기술에 대해 법정기한에 구애받지 않는 별도의 평가절차를 마련하여 검토하는 것이 고려될 수 있다. 하지만 이 절차도 동 평가결과의 활용도가 높지 않을 경우에는 추가적인 절차로 인해 업무

처리 지연 및 후단의 정책 반영률 저조에 대한 비판이 가능하므로 법적인 제도화가 필요할 것이다.

그동안 신의료기술평가는 지속적인 규제완화 정책으로 체계적문헌고찰을 수행하지 않는 신속문헌고찰을 통해 시장진입의 기회를 확대하였고, 건강보험 보장성 강화로 보험권 내에서의 합리적인 의료기술의 적정사용을 유도하기 위해서는 신의료기술평가 단계에서의 의료기술에 대한 합리적인 의사결정이 요구되는 실정에서, 이 연구는 기존의 평가체계를 바탕으로 개선안 마련을 위한 기반 틀을 제시하였다는 점에 의의가 있다. 본 연구를 바탕으로, 후속 연구를 통해 보다 안정적이고, 견고한 방법론으로 수행하여 신의료기술평가 권고 체계를 안착시키는 것이 필요할 것이다.

V

참고문헌

강현. How to understand and conduct evidence-based medicine. Korean Journal of Anesthesiology. 2016. 69(5): 435-445

김수경, 신상진, 박정수, 김준호, 백영지, 신채민, 김민정, 양장미. 의료기술평가제도의 국제 비교연구. 한국보건의료연구원. 2014.

김수영, 박지은, 서현주, 이윤재, 손희정, 장보형, 서혜선, 신채민. NECA 체계적 문헌고찰 매뉴얼. 한국보건의료연구원. 2011.

김수영, 최미영, 신승수, 지선미, 박지정, 유지혜, 류다현, 박승희. 임상진료지침 실무를 위한 핸드북. 한국보건의료연구원. 2015.

박동아, 설아람, 정유진, 최미영, 윤지은, 김윤정, 이나래, 박지정, 손수경, 한해림. 조건부 선별급여 재평가 연구. 2018. 건강보험심사평가원, 한국보건의료연구원

박동아, 이민, 김유림, 김윤정, 김지민, 김효정, 박은정, 박지정, 설아람, 유수연, 윤지은, 이나래, 이월숙, 이자연, 이종협, 이진이, 이현아, 조송희, 조애정, 최미영, 최인순, 최지은, 한경희, 한해림. 의료기술재평가체계 구축 및 운영방안 연구. 한국보건의료연구원. 2019.

박실비아, 박은자, 채수미, 이예슬. 신의료기술의 도입과 확산에 관한 정책분석. 한국보건사회연구원. 2014.

박은철, 이효정, 김주영, 남진영, 최재우, 전성연, 주영준, 현인선, 이남경. 국민 안전과 의료기술 조기활용을 위한 신의료기술평가 제도개선 방안 연구. 보건복지부, 연세대학교. 2016.

박종연. 의료기술의 평가와 급여결정 방향. 한국보건사회연구원. 2017.

서경화, 이선희, 신인순, 임선미, 장지은, 정유민, 박윤형. 근거 및 권고수준 등급화 체계 현황 분석. J Korean Med Assoc 2011;54(7):758-768.

최윤정, 채정미, 조수진, 최연미. 신의료기술평가 제도 성과평가 및 운영개선 방안 연구. 건강보험심사평가원. 2013

최윤정, 채정미, 최연미, 이보람. 가치에 근거한 권고등급 개발. 건강보험심사평가원 2014.

최미영, Guyatt GH. 근거기반의학의 과거, 현재 그리고 미래. Evidence and Values in Healthcare 2017 June 3(1): 1-5.

캐나다 보건부 홈페이지. URL:http://www.hc-sc.gc.ca/hcs-sss/medi-assur/index-eng.php

호주 보건부 홈페이지. URL: http://www.health.gov.au/internet/hta/publishing.nsf/content/common wealth-1

Alonso-Coello P, Schunemann HJ, Moberg J, Birgnardello-Petersen R, Akl EA, Davoli

M, Treweek S, Mustafa RA, Rada G, Rosenbaum S, Morelli A, Guyatt GH, Oxman AD, The GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016; 353.

Alonso-Colleo P, Oxman AD, moberg J. Birgnardello-Petersen R, Akl EA, Davoli M, Treweek S, Mustafa RA, Vandvik PO, Meerpohl J, Guyatt GH, Schunemann HJ, The GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016; 353. i2089.

Bipat S, van Leeuwen MS, IJzermans JN, Comans EF, Planting AS, Bossuyt PM, et al. Evidence-base guideline on management of colorectal liver metastases in the Netherlands. Neth J Med. 2007 Jan;65(1):5-14.

Dahm P, Oxman AD, Djulbegovic B, Guyatt GH, Murad MH, Amato L, Parmeli E, Davoli M, Moragn RL, Mustafa RA, Sultan S, Falck-Ytter Y, Akl EA, Schunemann HJ. Stakeholders apply the GRADE evidence-to-decision framework to facilitate coverage decisions. Journal of Clinical Epidemiology. 2017. 86 ; 129-139.

Guyatt G, Gutterman D, Baumann MH et al., Grading strength of recommendations and quality of evidence in clinical guidelines. CHEST 2006; 129: 174-81

Guyatt et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables, Journal of clinical epidemiology. 2011; 64(4): 383-394.

iEtD 홈페이지, Available from URL: https://ietd.epistemonikos.org/

Morgan RL, Kelley L, Guyatt GH, Johnson A, Lavis JN. Decision-making frameworks and considerations for informing coverage decisions for healthcare interventions: a critical interpretive synthesis.J Clin Epidemiol. 2018 Feb;94:143-150. doi: 10.1016/j.jclinepi.2017.09.023

Parmelli E, Amato L, Oxman AD, Pablo Anlnso-Coello, Bruentti M, Moberg J, Nonino F, Pregno S, Saitto C, Schunemann HJ, Davoli M, GRADE working group. GRADE EVIDENCE TO DECISION (EtD) FRAMEWORK FOR COVERAGE DECISIONS. International Journal of Technology Assessment in Health Care, 33:2 (2017), 176–182.

OECD Health Statistics 2018: Frequently Requested Data(2018, June)

SIGN 홈페이지. Available from URL: http://www.sign.ac.uk/

NECA 한국보건의료연구원
National Evidence-based Healthcare Collaborating Agency

발행일 2019. 10. 31.

발행인 이영성

발행처 한국보건의료연구원

이 책은 한국보건의료연구원에 소유권이 있습니다.
한국보건의료연구원의 승인 없이 상업적인 목적으로
사용하거나 판매할 수 없습니다.

ISBN : 978-89-6834-574-6