

임상진료지침 개발 매뉴얼 Ver 1.0

Guidance for development of clinical practice guidelines





Copyright © 한국보건의료연구원, 2011.

이 책의 저작권은 한국보건의료연구원에 있으며
무단 복제와 전제는 법으로 금지되어 있습니다.

임상진료지침 개발 매뉴얼

인쇄: 2011년 9월 30일 초판 1쇄 발행

발행인: 허대석

발행처: 한국보건의료연구원

서울시 종로구 원남동 28-7 창경빌딩

Tel. 02-2174-2700

www.neca.re.kr

ISBN 978-89-94581-64-4



다음과 같이 인용될 수 있습니다.

김수영, 지선미, 이수정, 이윤재, 박지은, 남미희, 박지애.
임상진료지침 개발 매뉴얼.
한국보건의료연구원. 2011.

김수영 교수
(대표저자) 한림대학교 의과대학 가정의학교실

지선미 책임연구원
 한국보건의료연구원

이수정 연구원
 한국보건의료연구원

이윤재 책임연구원
 한국보건의료연구원

박지은 연구원
 한국보건의료연구원

남미희 연구원
 한국보건의료연구원

박지애 연구원
 한국보건의료연구원

발간사

임상진료지침 개발의 필요성을 인식하고 국내에서 진료지침을 제작하고 보급한 이후 상당한 시간이 흘렀습니다. 하지만, 아직 한국의 의료 환경 속에서 진료지침을 어떻게 개발하는 것이 적절한지, 또 우리나라 의료제도 하에서 진료지침이 어떻게 활용되는 것이 바람직한지, 등에 대해서는 통일된 의견이 없었습니다.

진료지침 개발은 의학적(과학적) 근거에 기반하여 이루어져야 합니다. 따라서 우리나라에서 현재까지 개발된 진료지침의 제작과정에서 근거를 체계적으로 분석하고 정리했는지에 대한 검토가 필요합니다. 왜냐하면, 의료인들이 납득할 만한 근거가 제시되지 않은 진료지침은 진료현장에서 수용되기 어렵고, 보건의료정책에도 반영될 수 없기 때문입니다.

본 진료지침 개발 매뉴얼은 이러한 필요성을 감안하여, 한국의 진료지침 개발자들이 실무에 적용하기 쉽도록 진료지침 개발 순서에 따라 정리하였습니다. 또한 단계별로 ‘확인사항’을 제공하여 개발자들이 그 과정을 제대로 수행하였는지 직접 확인할 수 있게 하였습니다.

또, 진료지침이 정책에 반영되고 진료현장에 널리 수용되기 위해서는 사회적 가치 또한 반영되어야 합니다. 우리의 진료현장은 많은 이해당사자의 이해가 얽혀있는 곳으로, 특히 의료인들 사이의 의견일치는 사회적 합의를 유도하는 첫 출발점입니다. 따라서 진료지침을 개발하는 과정뿐만 아니라, 확산 및 보급하는 과정에서도 관련 분야의 전문가들의 합의를 유도해 낼 수 있는 전략을 고려해야 하는데 이 같은 접근도 이 매뉴얼에 담겨 있습니다.

앞으로도 한국보건의료연구원은 임상진료지침의 질 향상을 위해 지속적인 관심을 가지고 이를 위한 연구를 계속해 나갈 것입니다. 본 매뉴얼을 이용하여 근거중심의 임상진료지침이 많이 개발되고, 이를 통해 의료의 질이 향상되어 국민이 최적의 의료서비스를 받을 수 있기를 기대합니다.

2011년 9월

한국보건의료연구원 원장
허대석

머리말

임상 의사들은 매일 매일의 진료 현장에서 다양한 의사 결정을 해야 합니다. 이 환자를 정확하게 진단하려면 어떤 진단기기를 사용해야 할까? 이 환자에게 투약해야 할 약물은 무엇인가? 이 환자에게는 어떤 경우에 수술을 실시하여야 할까? 이런 의학적 의사 결정을 하는 것은 생각보다 쉽지 않습니다. 여러 가지 사항을 확인해야 하기 때문입니다. 해당 중재는 효과가 있는가? 효과에 대한 근거는 명확한가? 현재 이 환자의 성별, 연령, 동반 질환, 인종 등에 비추어 알려진 근거를 적용할 수 있을까?

임상진료지침은 이처럼 복잡한 의학적 의사 결정을 도와주기 위해 체계적인 방법으로 개발된 도구입니다. 만일 임상진료지침이 투명한 과정을 거쳐 현존하는 근거와 임상 의사의 경험, 환자 가치 등을 잘 결합하여 개발되었다면 임상진료지침은 의학적 의사 결정에 많은 도움을 주며 궁극적으로 보건의료의 질과 진료 결과의 향상에 이바지 할 것입니다.

우리나라에서 임상진료지침 개발의 역사는 결코 짧지 않습니다. 몇몇 연구자의 선구자적인 노력으로 80년대부터 일부 임상진료지침이 개발되기 시작하였지만 이들의 노력은 뿌리를 내리지 못하고 몇몇 학회에서 외국에서 개발된 진료지침을 번역하는 수준의 진료지침만을 출판하였을 뿐입니다. 이처럼 단편적인 움직임에 불과하던 임상진료지침은 몇몇 계기를 통해 변화의 움직임이 나타나기 시작하였습니다. 우선 최근 5-6년 동안 체계적 문헌고찰을 비롯한 근거중심의학적 방법론이 소개되기 시작하였고, 진료지침을 개발하는 것이 전문 학회의 학술 활동의 중요한 부분으로 인식되었으며, 정부, 보험기관 등 공적부분에서 임상진료지침의 개발과 활용이 의료의 질 향상, 합리적인 정책 결정에 매우 중요한 도구라는 사실을 깨닫기 시작하였습니다. 하지만 현재까지 우리나라에서 개발된 임상진료지침은 여러 가지 문제점을 가지고 있습니다. 우선 임상진료지침의 주제 선정 과정이 체계적이지 못하고 임상진료지침 개발그룹 구성원도 다학제성을 만족하지 못하고 방법론 전문가의 참여 정도도 미미합니다. 하지만 무엇보다도 중요한 문제점은 체계적 문헌고찰과 근거중심의학의 원리가 임상진료지침 개발 과정에 반영되지 않아서 투명성, 합리성, 신뢰성 등에서 만족스럽지 못한 모습을 보인다는 점입니다. 이러한 이유로 임상진료지침은 개발 이후 널리 보급, 실행되지 못하고 사장되는 결과를 낳고 있으며, 그러한 점에서 임상진료지침 개발과정에 대한 표준화된 방법론을 제시하는 것이 필요합니다.

한국보건 의료 연구원은 합리적이고 근거 중심적인 의사 결정에 도움을 주는 정보를 제공하는 미션을 가지고 있습니다. 그러한 점에서 합리적 의사 결정에 도움을 주는 임상진료지침은 한국 보건 의료 연구원의 중요한 연구 분야가 될 가능성이 있고 이는 외국의 주요 근거 중심 보건 의료 의사 결정 기관도 유사합니다. 한국 보건 의료 연구원에서는 이미 2009년 ‘임상진료지침 작성 간지 매뉴얼 0.5’를 발행하였지만 충분히 활용하거나 보급되지는 않았습니다. 2010년 NECA 체계적 문헌고찰 매뉴얼을 발행하면서 임상진료지침 개발 매뉴얼과 표준적 체계적 문헌고찰 방법론을 조율할 필요성이 제기되었고, 그런 필요성에 따라 이 매뉴얼이 기획되었습니다.

이 매뉴얼은 임상진료지침을 체계적 문헌고찰의 방법론에 따라 새롭게 개발하는 경우(*de novo development*)에 유용하며 기존에 발표된 ‘임상진료지침 수용개작 매뉴얼’과 같이 활용하면 더 많은 도움이 됩니다. 향후 임상진료지침 개발 매뉴얼과 수용개작 매뉴얼은 한국적 상황에 맞게 개정되고 통합되어 완결된 매뉴얼로 제시되어야 할 것입니다.

이 책을 만드는데 많은 수고를 하신 한국 보건 의료 연구원 지선미 연구원, 이수정 연구원, 박지은 연구원, 이윤재 연구원에게 감사드립니다. 아울러 책 발행에 물심양면 지원을 아끼지 않으신 허대석 원장님께서도 감사를 드립니다.

2011년 9월

대표저자
김수영

목차

14 요약

| 제 1부. 기획

- 20 1단계. 진료지침 주제 선정
- 26 2단계. 진료지침 개발그룹 구성
- 28 3단계. 기존 진료지침 검토
- 33 4단계. 개발 계획 수립
- 36 5단계. 핵심질문 결정

| 제 2부. 개발

- 38 6단계. 근거의 검색
- 44 7단계. 근거의 평가
- 54 8단계. 근거의 종합
- 57 9단계. 권고안 작성, 권고등급 결정
- 70 10단계. 합의안 도출

| 제 3부. 최종화

- 72 11단계. 외부검토 및 갱신 계획
- 74 12단계. 진료지침 출판

| 표 차례

23	표 1.	체계적 문헌고찰과 HTA 검색 자료원
24	표 2.	주요 기관 우선순위 선정 기준
25	표 3.	공식적, 비공식적 합의 도출 방법의 특징
29	표 4.	진료지침 검색 자료원
30	표 5.	진료지침 요약표의 예
31	표 6.	AGREE II 영역별 내용
45	표 7.	DAMI 이용자 매뉴얼
48	표 8.	비둘림 종류와 관련 영역
51	표 9.	연구 내 및 연구 간 비둘림 위험 요약 접근 방법
52	표 10.	비둘림의 종류와 평가 영역
52	표 11.	RoBANS 도구 평가영역
54	표 12.	근거표에 포함되는 항목
61	표 13.	근거수준의 의미
61	표 14.	권고강도 요인의 의미
67	표 15.	GRADE와 USPSTF, AHRQ의 근거수준 평가 방법
69	표 16.	체계적 문헌고찰 결론에 따른 추천되는 우리말 표현
74	표 17.	ADAPTE 진료지침 작성 양식
75	표 18.	임상진료지침 표준 보고가이드

목차

Ⅰ 그림 목차

- 14 그림 S. 임상진료지침 개발 과정
- 22 그림 1. PubMed에서 체계적 문헌고찰의 검색
- 22 그림 2. Cochrane Library에서 체계적 문헌고찰 검색
- 25 그림 3. 주제 확인, 주제 선정, 주제 가공의 과정
- 39 그림 4. COSI 모델에서 제안하는 검색원 범위
- 41 그림 5. EMBASE의 Emtree 예시
- 42 그림 6. 검색과정(SIGN)
- 42 그림 7. 문헌선정 과정
- 43 그림 8. 문헌 선정 흐름도(PRISMA flow chart)
- 45 그림 9. 문헌의 분류도구, DAMI
- 49 그림 10. 비풀림 위험 평가 결과 예시(전체 결과)
- 50 그림 11. 비풀림 위험 평가 결과 예시(연구별)
- 55 그림 12. GRADEpro를 이용하여 만든 근거 요약표
- 56 그림 13. GRADEpro에서 만든 근거 요약표를 RevMan으로 편입한 결과
- 59 그림 14. GRADE 과정
- 60 그림 15. Evidence profile
- 62 그림 16. GRADE 과정 흐름도

Ⅰ 부록

- 78 부록1. NECA 이해상충 선언(COI) 양식
- 79 부록2. 체계적 문헌고찰 질 평가도구, AMSTAR
- 83 부록3. 무작위배정연구 질 평가도구, Cochrane ROB
- 92 부록4. 비무작위 연구의 질 평가도구, RoBANS
- 99 부록5. 임상진료지침 표준 보고가이드, STARIGs

1. 진료지침 개요

- 진료지침은 “보건의료 제공자의 의사결정에 도움을 주기 위해 과학적 근거를 검토하여 체계적으로 개발된 권고와 관련 내용을 기술한 것”으로 정의할 수 있다.
- 종설(narrative review), 체계적 문헌고찰(systematic review), 환자용 지침만 있는 경우, 교육매뉴얼, 단순 번역된 진료지침, 의료 기술 평가(health technology assessment, HTA), 진료 계획표(critical pathway) 등은 진료지침과는 구별된다.

2. 진료지침 개발 과정

- 진료지침 개발 과정은 다음과 같이 3부, 12단계로 구분한다.

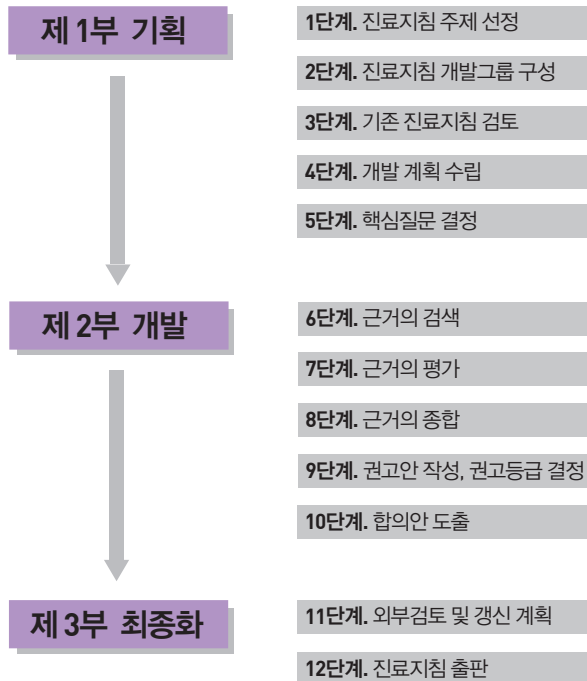


그림 5. 임상진료지침 개발 과정

제 1부. 기획

1단계. 진료지침 주제 선정

- 진료지침의 주제 선정은 체계적인 과정을 거쳐야 한다.
- 진료지침의 주제 선정은 현안이 되고 있는 주제에 대한 확인(topic identification), 적절한 우선 순위 원칙에 따른 주제 선택(topic selection), 주제를 기획에 맞도록 구체화하는 주제 가공(topic refinement)의 3단계를 거친다.
- 주제 확인은 주요 현안 확인, 주제 제안, 진료지침 검토, 관련 체계적 문헌고찰 검토, 관련 HTA 확인 등의 방법을 이용한다.
- 우선순위 선정 기준을 마련하고 공식적 합의 절차를 거쳐 진료지침 주제를 선정한다. 비용, 질병부담, 영향 잠재력, 사회적 관심사, 새로운 근거, 충분한 근거, 진료의 변이 등이 선정기준에 쓰일 수 있다.
- 주제 확인과 주제 선정 절차를 거쳐 확정된 주제에 대하여 가공을 통해 구체화 작업을 한다. 만일 주제가 진료지침보다는 체계적 문헌고찰 등 다른 형태의 연구를 해야 하거나 임상연구를 기획해야 하는 경우는 이에 맞는 절차를 거쳐야 한다.

2단계. 진료지침 개발그룹 구성

- 진료지침 개발 실무를 담당할 개발그룹과 운영위원회 등의 관련 조직을 구성한다.
- 개발그룹 구성 시에는 다학제성, 대표성, 지속성을 고려한다.
- 개발그룹 혹은 운영위원회는 운영약관 마련, 이해상충 선언, 합의원칙 결정, 잠재적 승인 기구 선정, 저자됨 원칙 결정 등 기획업무를 수행한다.

3단계. 기존 진료지침 검토

- 기존 진료지침 검토 과정은 진료지침 검색과 진료지침 요약표 작성, 평가의 과정을 거친다.
- 진료지침을 검색할 때는 NGC, G-I-N, PubMed, 그리고 국내 데이터베이스(KGC, KOMGI)는 반드시 확인해야 한다.
- 진료지침 선택을 위한 포함·제외기준을 결정한다.
- AGREE II를 이용하여 진료지침의 질을 평가한다.

4단계. 개발 계획 수립

- 진료지침의 범위와 목적을 PIPOH (Population, Intervention, Professionals/Patients, Outcomes, Healthcare settings)에 따라 결정한다.
- 진료지침의 존재 여부, 수용개작 방법론에 대한 신뢰 등을 고려하여 신규 개발 또는 수용개작할 지 여부를 결정한다.
- 진료지침의 실행과정, 업무, 수행주체, 시간표 등을 명시한 진료지침 개발 계획서를 작성한다.

5단계. 핵심질문 결정

- 핵심질문을 PICO (Patient/Problem, Intervention, Comparison, Outcomes)의 4가지 요소를 포함하여 구체적인 질문으로 만든다.
- 핵심질문의 목록을 진료지침에 제시한다.

제 2부. 개발

6단계. 근거의 검색

- 검색의 민감도를 최대한으로 높인 포괄적인 검색을 실시한다.
- 검색의 범위에 적절한 데이터베이스를 선택한다.
- 진료지침의 범위에 적절한 검색어로 검색전략을 수립한다.
- 체계적 문헌고찰, 무작위 대조 연구 등의 순서로 문헌검색을 진행한다.

7단계. 근거의 평가

- 근거 평가는 연구설계 분류, 연구설계에 적합한 평가 도구 결정, 근거 평가 순으로 진행된다.
- 연구설계 분류는 study Designs Algorithm to assess the effectiveness of Medical Interventions (DAMI)를 이용해서 시행한다.
- 연구설계에 적합한 질 평가 도구를 이용해서 문헌의 질을 평가한다.
- 체계적 문헌고찰은 Assessment of Multiple SysTemAtic Reviews (AMSTAR), 무작위배정 비교 임상시험은 코크란의 Risk Of Bias (ROB) 도구, 비무작위연구는 Risk of Bias Assessment tool for Non-randomised Study (RoBANS)를 이용하여 평가한다.

8단계. 근거의 종합

- 근거의 종합을 위해 개별 연구에 대해서 연구 특성, 연구 대상자, 중재, 결과, 문헌의 질 평가 등을 내용으로 하는 근거표를 작성한다.
- 주요 결과 각각에 따라 전반적 근거를 요약한 근거 요약표(summary of findings table)를 작성할 수 있다.

9단계. 권고안 작성, 권고등급 결정

- 근거수준은 현재까지의 근거로 해당 중재의 효과에 대해 확신하는 정도이다. 근거수준 평가는 GRADE에 준해서 시행하되 분야, 목적 등에 맞게 수정되어 사용될 수 있다.
- 근거수준은 ① 비뿔립 위험, ② 이질성(일관성 없음), ③ 비직접성, ④ 비정밀, ⑤ 출판 비뿔림이 있는 경우 근거수준 정도를 1단계 혹은 2단계를 낮추고, 만일 ① 효과의 크기가 큼, ② 교란변수의 영향, ③ 양-반응 관계가 있으면 단계를 높인다. 최종적으로 '높음', '중등도', '낮음', '불충분'의 네 단계로 근거수준을 평가한다.
- 권고등급은 권고 대상 환자에서 중재를 시행하였을 때 원하는 결과가 원하지 않는 결과보다 많이 발생할 것으로(혹은 적게 발생할 것으로) 확신하는 정도를 말한다. 권고 등급은 일반적으로 ① 근거수준, ② 이득과 위해의 저울질, ③ 가치와 선호도, ④ 자원 이용(비용)의 네 가지를 고려하여 결정한다.
- 권고등급은 strong for: A, weak for: B, weak against: C, strong against: D, insufficient: I로 분류한다.
- 권고안은 특이적이며 모호하지 않아야 한다. 주요 권고안은 쉽게 확인할 수 있어야 하며, 권고와 이를 지지하는 근거들이 명시적으로 연결되어 있어야 한다. 또한 권고 등급이 제시되어야 한다.

10단계. 합의안 도출

- 권고안 도출과정에서 이용된 합의 도출 방법을 기술한다.
- 권고안 내용에 이견이 적을 경우 비공식적 합의 방법론으로 권고안을 도출할 수 있다.
- 권고안 내용에 이견이 있거나 근거가 부족할 경우 공식적 합의 방법론을 이용하여 권고안을 도출하는 것이 좋다.

제 2부. 개발

11단계. 외부검토 및 갱신 계획

- 진료지침 초안이 완성되면 해당분야 전문가나 이해당사자에게 초안을 보내고 그에 대한 의견을 받을 수 있다.
- 외부검토는 동료검토(peer review), 이해당사자의 의견 수렴, 사용자 사전 조사로 구분할 수 있다.
- 갱신 계획을 세운다.
- 진료지침 주제와 관련된 단체의 공식적 승인을 받는다.

12단계. 진료지침 출판

- 실제 사용될 진료 상황에 맞게 진료지침을 맞춤화하여 출판한다.
- 진료지침 보고지침 양식을 준용한다.

제 1부. 기획

1단계. 진료지침 주제 선정

요약

- 진료지침의 주제 선정은 체계적인 과정을 거쳐야 한다.
- 진료지침의 주제 선정은 현안이 되고 있는 주제에 대한 확인(topic identification), 적절한 우선순위 원칙에 따른 주제 선택(topic selection), 주제를 기획에 맞도록 구체화하는 주제 가공(topic refinement)의 3단계를 거친다.
- 주제 확인은 주요 현안 확인, 주제 제안, 진료지침 검토, 관련 체계적 문헌고찰 검토, 관련 HTA 확인 등의 방법을 이용한다.
- 우선순위 선정 기준을 마련하고 공식적 합의 절차를 거쳐 진료지침 주제를 선정한다. 비용, 질병부담, 영향 잠재력, 사회적 관심사, 새로운 근거, 충분한 근거, 진료의 변이 등이 선정기준에 쓰일 수 있다.
- 주제 확인과 주제 선정 절차를 거쳐 확정된 주제에 대하여 가공을 통해 구체화 작업을 한다. 만일 주제가 진료지침보다는 체계적 문헌고찰 등 다른 형태의 연구를 해야 하거나 임상연구를 기획해야 하는 경우는 이에 맞는 절차를 거쳐야 한다.

1. 주제 선정 절차

진료지침 개발에 있어 적절한 주제를 선정하는 것은 매우 중요하다. 진료지침은 부담이 높고 임상적으로 중요한 질병의 관련 주제 중 근거와 진료 사이에 간극(gap)이 있는 부분에 대하여 적절한 해답을 주어야 한다. 이러한 과정은 적절한 주제 선정을 통해 구체화된다. 진료지침의 주제 선정은 현안이 되고 있는 주제에 대한 확인(topic identification), 적절한 우선순위 원칙에 따른 주제 선택(topic selection), 주제를 기획에 맞도록 구체화하는 주제 가공(topic refinement)의 3단계를 거친다.

1) 주제 확인(Topic Identification)

현재 진료지침의 개발이 필요한 주제는 무엇인지에 대해서 확인하는 단계이다. 다양한 경로를 통해 필요한 주제가 무엇인지를 확인한다. 일반적으로 주요 현안 확인, 주제 제안, 진료지침 검토, 관련 체계적 문헌고찰 검토, 관련 HTA 확인 등의 방법을 이용한다.

(1) 주요 현안 확인

보건의료 분야에서 진료지침 개발이 필요한 주제가 무엇인지를 확인하는 것으로 국민의 주요 관심사를 확인한다. 이러한 주제는 언론이나 방송에서 이슈가 되는 사안일 수도 있고, 많은 사람들이 관심을 가질 만한 주제일 수도 있다.

(2) 주제 제안

NECA 주제 공모, 임상연구센터별 공모, 학회별 공모, 주요 연구자에게 시행하는 공모 등을 통해 진료지침 개발이 필요한 주제를 확인할 수 있다.

(3) 진료지침 검토

해당 질환의 진료지침을 검색하여 검토하는 것은 주제 확인의 대표적인 방법이다. 진료지침을 검토하면 현재 쟁점인 부분이 무엇인지 파악할 수 있고 국가별로 어떤 내용을 중요하게 생각하는 지에 대해서 알 수 있게 해줄 뿐만 아니라 핵심질문의 전체 리스트를 확인할 수 있어서 유용하다.

진료지침 검토는 진료지침의 검색, 진료지침의 선정, 진료지침의 비교 등으로 이루어지며 이 부분은 제 3단계. 기존 진료지침의 검토를 참조한다.

(4) 기존 체계적 문헌고찰, 관련 HTA (health technology assessment) 확인

해당 질환에 대해서 시행된 기존 체계적 문헌고찰을 검색하여 검토하면 해당 질환의 주요 핵심질문을 확인할 수 있고 현재까지의 근거를 확인할 수 있다.

체계적 문헌고찰은 PubMed에서 질환 관련 키워드와 함께 출판 형태로 "meta-analysis"를 넣거나(그림 1), 메뉴 "clinical queries"에서 "find systematic reviews"를 검색한다. Cochrane Library에서 검색 후 "Cochrane reviews"와 "DARE"에 있는 목록을 확인한다(그림 2). HTA 목록을 확인하면 각국의 주요 현안을 확인하고 국내의 잠재적 현안을 확인할 수 있는 장점이 있다. HTA 목록은 주요 HTA site에서 해당 질환에 대한 검색을 시행하며 HTA site는 표 1과 같다.



그림 1. PubMed에서 체계적 문헌고찰의 검색



그림 2. The Cochrane Library에서 체계적 문헌고찰 검색

1단계.
진료지침
주제선정

2단계.
진료지침
개발그룹
구성

3단계.
기존
진료지침
검토

4단계.
개발
계획수립

5단계.
핵심질문
결정

표 1. 체계적 문헌고찰과 HTA 검색 자료원

관련기관	URL
National Institute for Health and Clinical Excellence (NICE)	http://www.nice.org.uk
Canadian Agency for Drugs and Technologies in Health	http://www.cadth.ca
Canadian Medical Association Infobase	http://mdm.ca/cpgsnew/cpgs/index.asp
The Cochrane library	http://www.thecochranelibrary.com
Food and Drug Administration	http://www.fda.gov
Centre for Reviews and Dissemination Health Technology Assessment Database	http://www.york.ac.uk
Haute Autorité de Santé (HAS)	http://www.has-sante.fr
CHU de Rouen - Catalogue & Index des Sites Médicaux Francophones (CISMef)	http://doccismef.chu-rouen.fr/servlets/Simple?Mot=recommandations+professionnelles&aff=4&tri=50&datt=1&debut=0&rechercher.x=29&rechercher.y=18
Bibliothèque médicale AF Lemanissier	http://www.bmlweb.org/consensus.html
Direction de la lutte contre le cancer - Ministère de la santé et des services sociaux du Québec	http://mss.gouv.qc.ca
SOR : Standards, Options et Recommandations	http://www.fnclcc.fr
Registered Nurses Association of Ontario	http://www.rnao.org
Agency for Quality in Medicine	http://www.aeqz.de

2) 주제 선택: 우선순위 선정

주요 현안 확인, 주제 제안, 진료지침 검토, 관련 체계적 문헌고찰 검토, 관련 HTA 확인 등을 통해 수집된 주제에 대하여 적절한 절차에 따라서 우선순위를 선정한다. 확인된 주제 중 정보가 충분하지 않은 경우는 간단한 문헌고찰을 통해 주제 선택에 필요한 정보를 확보해야 한다.

(1) 선정 기준

다양한 기준으로 우선순위를 선정할 수 있다.

- 유효률
- 중재의 과다사용, 과소사용, 오남용의 존재
- 질병 부담
- 진료변이 여부
- 치료방법에 따른 비용 차이
- 진료자침이 실제 진료에 영향을 미칠 가능성
- 진료자침이 진료의 질과 결과에- 생존율, 삶의 질을 향상시킬 잠재력
- 수용개선에 사용할 수 있는 양질의 근거중심 진료자침의 존재 여부

주요 기관에서 사용하고 있는 우선순위 선정 기준은 아래와 같다.

표 2. 주요 기관 우선순위 선정 기준

기관	비용	질병부담	영향 잠재력	사회적 관심사	새로운 근거	충분한 근거	진료의 변이
AHRQ	V	V	V	V		V	V
BCBSA TEC			V	V	V		
CADTH	V	V	V	V		V	
MedCAC and CMS			V	V	V		
NICE	V	V	V			V	
NIH OMAR	V	V	V	V	V	V	
DERP	V						
USPSTF		V	V	V	V	V	V

*CADTH = Canadian Agency for Drugs and Technologies in Health; DERP = Drug Effectiveness Review Project; NICE = National Institute for Health and Clinical Excellence; NIH OMAR = National Institutes of Health Office of Medical Applications of Research.

(2) 선정 방법

위에서 설정한 우선순위 기준에 따라 공식적 합의 방법론에 의해서 주제를 선정하는 것이 일반적이다. 공식적 합의 방법론은 비공식적 합의 방법론의 문제점인 한 두 사람에 의한 회의 주도, 여러 대안에 대한 검토 부족 등을 해결하기 위해서 제시된 방법론이다. 공식적 합의 방법론은 결정의 익명성, 과정의 통제화, 과학적 신뢰성 등을 특징으로 하며 현재 주로 사용되고 있는 공식적 합의 도출 방법은 델파이법, 명목 집단 기법(Nominal group technique, NGT), 합의 도출 회의 세 가지이다. 이들 방법론은 ① 우편 설문지를 쓰는가?, ② 각 개인이 독립적인 결정을 '개별적으로' 하는가, ③ 비밀보장은 어느 정도 이루어지는가?, ④ 각 집단간 대면 접촉이

이루어지는가, 만일 그렇다면 그것은 구조적인가?, ⑤ 각 참여자의 견해를 어떻게 종합하는가? 등으로 구별된다(표 3).

표 3. 공식적, 비공식적 합의 도출 방법의 특징

방법	우편 설문	개인적 결정	집단 선택에 대한 공식적 피드백	대면 접촉	구조화된 상호작용	결합 방법
비공식	X	X	X	0	X	함축적
델파이법	0	0	0	X	0	명시적
NGT	X	0	0	0	0	명시적
RAND법	0	0	0	0	0	명시적
합의 도출 회의	X	X	X	0	X	함축적

3) 주제 가공

주제 확인과 주제 선정 절차를 거쳐 확정된 주제에 대하여 가공을 통해 구체화 작업을 한다. 만일 주제가 진료지침보다는 체계적 문헌고찰 등 다른 형태의 연구를 해야 하거나 임상연구를 기획해야 하는 경우는 이에 맞는 절차를 거쳐야 한다. 이를 요약하면 그림 3과 같다.

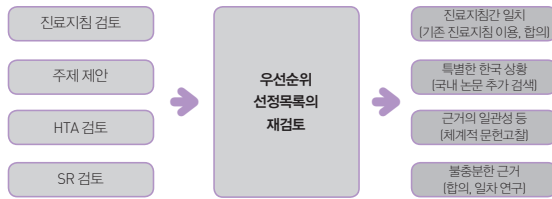


그림 3. 주제 확인, 주제 선정, 주제 가공의 과정

2. 주제 선정 기획 회의

기구 혹은 센터 단위로 진료지침 주제 선정을 위한 임시 혹은 영구 조직을 운영한다. 주제 선정 관련 기획은 장기 기획과 단기 기획이 모두 필요하며, 장기 기획은 3-4년 단위로 하고 단기 기획은 1년 단위로 시행할 수 있다. 장기 기획은 “진료지침 전략 기획 회의”의 형태로 이루어진다. 진료지침 관련 계획 수립, 큰 그림에서의 사업 수행, 임상연구와 연계, 사업의 성격 파악 등이 이루어진다. 이 때 주제 확인, 우선순위 확인, 주제 결정 및 세부 조정의 모든 과정을 다 거치게 된다. 사업 초기 혹은 3년 정도의 주기로 실시할 수 있다. 단기 기획은 전체적 계획 속에서 필요한 진료지침의 주제를 선정하며, 해마다 현안과 근거의 변화를 반영하고, 방법론적 보완, 내용의 수정이 필요한 지 검토하게 된다.

제 1부. 기획

2단계. 진료지침 개발그룹 구성

요약

- 진료지침 개발 실무를 담당할 개발그룹과 운영위원회 등의 관련 조직을 구성한다.
- 개발그룹 구성 시에는 다학제성, 대표성, 지속성을 고려한다.
- 개발그룹 혹은 운영위원회는 운영약관 마련, 이해상충 선언, 합의원칙 결정, 잠재적 승인 기구 선정, 저자됨 원칙 결정 등 기획업무를 수행한다.

진료지침 개발그룹(이하 개발그룹)은 실무위원회라고도 하며, 진료지침 개발 계획부터 진료지침 초안 작성, 외부검토 후 최종 진료지침 출판을 책임진다. 진료지침의 신뢰성은 개발그룹의 신뢰성에 달려 있으므로 개발그룹에 포함된 사람이 누구인가와 그들의 전문지식과 기술이 무엇인가가 중요하다. 개발그룹은 진료지침의 범위에 따라서 10-20명 정도로 구성된다. 또한 진료지침 개발 과정을 잘 이해하기 위해 개발그룹을 위한 교육이 필요하다.

개발그룹 외에도 운영위원회, 자문위원회, 검토위원회 등을 필요에 따라 구성하기도 한다. 진료지침 개발 관련 조직은 진료지침의 성격, 범위, 분량 등에 따라서 적절히 선택하여 구성한다. 운영위원회는 과정 전체를 총괄하여 진료지침 범위 결정, 개발방법론 결정, 조직구성(예: 개발그룹 구성), 운영약관 제정, 개발계획 수립 등을 책임질 수 있다. 자문위원회는 진료지침 개발 각 단계에 맞는 자문을 하기도 하고, 진료지침 개발 후 동료검토를 수행할 수도 있다. 각 위원회의 구성원은 완전히 분리될 수도 있고, 일부만 포함될 수도 있으며, 합동 회의를 가질 수도 있다.

개발그룹 구성 시에는 다음의 사항을 고려하여 구성한다.

• 다학제성

관련된 모든 분야 임상 전문가와 방법론 전문가, 검색 전문가 등으로 다학제적으로 구성한다. 개발그룹에는 주요 이해당사자(key stakeholders) 대부분이 포함되어야 하며, 최소한 두 분야 이상의 임상 전문가와 방법론 전문가는 반드시 포함되어야 한다.

• 대표성

진료지침의 실행과 확산을 고려하여, 개발에 참여하는 학회 또는 기구를 대표할 수 있어야 한다.

• 지속성

진료지침 개발 이후 실행, 확산 및 갱신 등을 지속적으로 수행할 수 있는 조직으로 구성해야 한다.

개발그룹은 다음 항목들을 마련하거나 고려해야한다.

- **운영약관(terms of reference):** 운영위원회나 개발그룹에서 작성하며, 업무의 범위, 위원 구성, 회의 주기 등을 포함한다.
- **이해상충 선언:** 진료지침 개발에 관련된 모든 위원들은 NECA 이해상충 공개서를 작성하고 서명해야 한다. 이해상충 문제가 있다면 이를 어떻게 다룰 것인가에 대해서도 명확히 규정해야 한다(부록 1).
- **합의 과정:** 운영위원회나 개발그룹에서 어떻게 합의에 이를 것인가(예: 공식적·비공식적인 합의과정), 그리고 합의 과정을 어떻게 보고할 것인가를 결정해야 한다.
- **잠재적 승인기구 선정:** 개발된 진료지침을 공식적으로 승인(endorsement) 받을 기구 또는 단체를 선정해야 한다. 만약 필요하다면, 승인기구 대표자를 개발그룹 혹은 운영위원회 구성원으로 개발과정에 참여시키거나, 진료지침의 외부 검토 과정에 참여시킬 것을 고려해야 한다.
- **저자됨(authorship):** 진료지침의 초안은 누가 작성하고, 최종보고서는 누가 작성할 지 그리고 저자됨의 원칙, 저자의 순서, 그룹 저자 여부 등을 결정한다.
- **실행 및 확산 전략:** 웹사이트에 발표, 학술지에 투고 등 확산 전략을 결정한다. 실행전략을 준비 단계부터 고려한다.

확인사항	예	아니오
1. 개발그룹은 다학제적(최소한 2개 분야 임상 전문가와 방법론 전문가)으로 구성되었다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 진료지침 개발 구성원들의 이해상충 문제를 확인하였다.	<input type="checkbox"/>	<input type="checkbox"/>
3. 개발그룹에 대한 상세 정보가 기술되어 있다.	<input type="checkbox"/>	<input type="checkbox"/>
4. 합의 원칙을 결정하였다.	<input type="checkbox"/>	<input type="checkbox"/>

제 1부. 기획

3단계. 기존 진료지침 검토

요약

- 기존 진료지침 검토 과정은 진료지침 검색과 진료지침 요약표 작성, 평가의 과정을 거친다.
- 진료지침을 검색할 때는 NGC, G-I-N, PubMed, 그리고 국내 데이터베이스(KGC, KOMGI)는 반드시 확인해야 한다.
- 진료지침 선택을 위한 포함·제외기준을 결정한다.
- AGREE II를 이용하여 진료지침의 질을 평가한다.

기존 진료지침 검토는 현재 개발된 진료지침의 현황을 파악하는 것이다. 기존 진료지침 검토의 목적은 첫째, 이미 진료지침이 개발되어 있는 경우 중복된 작업을 피하는 것이고 둘째, 진료지침을 새로 개발할 지 수용개발할 지 의사 결정에 도움을 주며 셋째, 개발계획 수립을 위한 기본적 방향설정을 위한 정보를 얻기 위함이다.

1단계에서 언급한 것과 같이 기존 진료지침 검토는 진료지침 주제 선정에도 중요한 역할을 한다.

기존 진료지침 검토 과정은 진료지침 검색, 진료지침 선정, 진료지침 요약표 작성, 진료지침 평가의 과정을 거친다.

1. 진료지침 검색

진료지침 관련 홈페이지, 국내외 주요 전자 데이터베이스를 이용한다. 진료지침은 학술지에 실리지 않을 수 있고, 문헌 데이터베이스에 없을 수 있기 때문에 검색은 미국의 National Guideline Clearinghouse (<http://www.guideline.gov>), Guidelines International Network (<http://www.g-i-n.net>)과 같은 진료지침 정보센터나 특정 국가 데이터베이스에서 시작하는 것이 좋다 (표4).

PubMed 검색을 통해 추가적인 진료지침 목록을 찾을 수 있다. 이때 진료지침 관련 검색어 ("guideline[Publication Type] OR practice guideline[Publication Type] OR recommendation*[Title] OR standard*[Title] OR guideline*[Title]")와 임상 주제와 관련된 검색어를 결합하여 사용한다.

1단계. 진료지침 주제선정

2단계. 진료지침 개발그룹 구성

3단계. 기존 진료지침 검토

4단계. 개발 계획수립

5단계. 핵심질문 결정

구글, 야후 등과 같은 인터넷 검색 엔진 역시 진료지침을 찾는 데 사용할 수 있다. 특히 인터넷 검색은 일본어, 중국어 진료지침을 확인하는데 유용하다. 국내 진료지침은 Korean Guideline clearinghouse (KGC, <http://www.cpg.or.kr>), 진료지침정보센터 (KoMGI, <http://www.guideline.or.kr>)를 검색한다. 또한 KoreaMed (<http://www.koreamed.org>), Medric (<http://www.medic.or.kr>) 등에서 국내 진료지침에 대한 검색을 할 수 있다.



표 4. 진료지침 검색 자료원

관련 기관	URL
국외	
Guidelines International Network (G-I-N)	http://www.g-i-n.net/
National Guidelines Clearinghouse (NGC)	http://www.guideline.gov/
Ontario Guidelines Advisory Committee (GAC) Recommended Clinical Practice Guidelines	http://www.gacguidelines.ca
Institute for Clinical Systems Improvement (ICSI)	http://www.icsi.org
National Institute for Health and Clinical Excellence (NICE)	http://www.nice.org.uk
New Zealand Guidelines Group	http://www.nzgg.org.nz
Scottish Intercollegiate Guidelines Network (SIGN)	http://www.sign.ac.uk/guidelines/index.html
Haute Autorité de Santé (HAS)	http://www.has-sante.fr
Finnish Medical Society Duodecim	http://www.kaypahoito.fi
American Society of Clinical Oncology	http://www.asco.org
Cancer Care Ontario Practice Guideline Initiative	http://cancercares.on.ca
National Cancer Institute	http://www.cancer.gov
National Comprehensive Cancer Network	http://www.nccn.org
국내	
KGC (Korean Guideline Clearinghouse)	http://www.cpg.go.kr
진료지침정보센터(KoMGI)	http://www.guideline.or.kr

2. 진료지침 선택을 위한 포함·제외 기준 결정

운영위원회나 개발그룹은 검색된 진료지침의 일차선택에 이용할 포함·제외 기준을 미리 정해야 한다.

다음과 같은 기준을 적용할 수 있다.

1) 선택기준

- 근거중심 진료지침(체계적 문헌 검색에 대한 보고가 있고, 권고와 지지 근거 사이에 명확한 연계가 있는 것)만 선택
- 국가 단위 또는 국제적 단위 진료지침만 선택
- 특정 기간에 발표된 진료지침만 선택
- 중요한 체계적 문헌고찰이 출간된 이후에 개발된 것만 선택
- 동료검토(peer review)가 이루어진 진료지침만 선택
- 특정언어로 작성된 진료지침만 선택

2) 배제기준

- 대표성 없는 단일저자 진료지침 배제
- 참고문헌없이 출판된 진료지침 배제

3. 진료지침 요약표 작성

검색된 진료지침의 특성에 대한 요약표를 작성한다. 요약표에 핵심질문별 포함 여부, 진료지침의 범위 등을 넣을 수도 있다(표 5).

표 5. 진료지침 요약표의 예

구분	진료지침1	진료지침2	진료지침3
제목			
개발자			
국가			
출판연도			
핵심질문 1			
핵심질문 2			
:			
진료지침 범위			

4. 진료지침의 평가

진료지침의 질 평가는 기획 단계에서 할 수도 있고, 수용개작(adaptation)을 하는 경우라면 진료지침의 범위와 목적을 결정한 이후에 할 수도 있다. 진료지침의 질을 평가하는 도구로 AGREE II (Appraisal of Guidelines Research and Evaluation)와 국내에서 개발 예정인 한국형 진료지침 평가 도구(Quality Assessment of Adapted Guidelines, QWADAG) 등이 있다.

1) AGREE II

AGREE II는 진료지침의 질을 평가하기 위해 만든 도구로 주로 진료지침의 개발과정을 평가한다. 이 도구는 범위와 목적, 이해당사자의 참여 등 6개 영역에 23개 문항으로 구성되어 있다(표 6).

표 6. AGREE II 영역별 내용

영역	문항번호	내용
범위와 목적	1-3	전반적인 목적, 특정 임상질문, 대상 인구집단에 대한 내용
이해당사자의 참여	4-7	진료지침 사용자의 견해를 어느 정도 대표하고 있는가에 대한 내용
개발의 엄격성	8-14	근거를 수집하고 결합하는 과정과 권고를 만들어 내고 갱신하는 방법에 대한 내용
명확성과 표현	15-17	진료지침의 언어와 형식에 대한 내용
적용성	18-21	진료지침을 적용할 때 발생할 수 있는 조직적, 행동적, 비용적 영향에 대한 내용
편집의 독립성	22-23	권고의 독립성, 진료지침 개발그룹에서 생길 수 있는 이해상충에 대한 내용

AGREE II 도구는 최소한 2명이 평가하며, 평가의 신뢰성을 위해 4명이 평가하는 것을 추천한다. AGREE II의 개별문항은 7점 척도로 평가하며, 각 영역의 개별 문항 점수를 합산하여 해당 영역의 최대 점수에 대한 백분율로 표준화한다. 모든 영역을 평가한 후에 개별 진료지침의 전반적인 평가를 7점 척도로 평가하며, 진료지침의 권고여부에 대해서 '권고함', '변경을 전제로 권고함', '권고하지 않음' 중 하나를 선택한다.

2) 수용개작 진료지침 평가도구

K-GINE은 진료지침의 질을 평가하는 도구로 외국 근거의 국내 적합성 및 국내 연구가 필요한 부분에 대한 검토 등과 같은 항목을 포함하고 있으며, 2009년 국내에서 개발되었다. 또한 수용개작된 진료지침 질 평가를 위한 도구가 2011년에 개발될 예정이다.

확인사항	예	아니오
1. 진료지침을 포괄적으로 검색하였다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 진료지침 선정 원칙을 정하였다.	<input type="checkbox"/>	<input type="checkbox"/>
3. 기존 진료지침의 권고 내용을 검토하고 비교하였다.	<input type="checkbox"/>	<input type="checkbox"/>
4. 검토한 기존 진료지침의 질을 평가하였다.	<input type="checkbox"/>	<input type="checkbox"/>

4단계. 개발 계획 수립

요약

- 진료지침의 범위와 목적을 PIPOH에 따라 결정한다.
- 진료지침의 존재 여부, 수용개작 방법론에 대한 신뢰 등을 고려하여 신규 개발 또는 수용개작할 지 여부를 결정한다.
- 진료지침의 실행과정, 업무, 수행주체, 시간표 등을 명시한 진료지침 개발 계획서를 작성한다.

1. 진료지침 범위와 목적의 결정

진료지침의 목적을 구체화시키고 진료지침이 다룰 범위를 명백하게 결정한다. 진료지침의 목적은 진료지침이 사회나 환자 집단에 미칠 잠재적 영향에 대한 것으로 구체적으로 작성한다. 예를 들어 '본 진료지침은 비용-효과적인 방법으로 항우울제를 합리적으로 처방하는 것을 목적으로 한다.'와 같이 작성한다.

진료지침의 범위는 개발비용 및 기간, 질병부담 및 임상적 중요성, 근거의 질과 양 등을 고려하여 설정한다. 이는 PIPOH (Population, Intervention, Professional, Outcomes, Healthcare settings) 도구를 이용하여 구체화할 수 있다.

- **인구집단(Population)** : 대상 인구집단의 특성. 나이, 성별, 임상적 특성, 동반 질환 등
- **중재(Intervention)** : 논의할 중재
- **전문가/환자(Professionals/Patients)** : 진료지침의 목표 사용자로서 전문가 혹은 환자
- **결과(Outcomes)** : 환자 결과(예: 생존율 증가나 삶의 질 향상), 시스템 결과(예: 임상 변이 감소), 공중보건 결과(예: 자궁경부암 유병률 감소)
- **의료 환경(Healthcare settings)** : 진료지침이 실제 쓰일 의료 환경(예: 1차, 2차, 3차 진료 등)과 지역 상황

1단계.

진료지침
주제선정

2단계.

진료지침
개발그룹
구성

3단계.

기존
진료지침
검토

4단계.

개발
계획수립

5단계.

핵심질문
결정

예시) 캐나다 자궁암 선별 진료지침

- **Population** : 중등도 위험 여성만 다루기로 하였다(즉 HIV 양성, 최근 5년 이내에 자궁도말 세포진 검사에서 중등도의 이형성(dysplasia)이 있었던 여성은 제외함). 선별의 시작 연령이나 종결 연령은 다루지 않기로 하였다.
- **Intervention** : 선별 중재만을 다루기로 하였다. 특정 장비(즉 통상적인 자궁경부 세포진검사 혹은 liquid based cytology)로 진료지침 검색 범위를 제한하지 않았다.
- **Professionals** : 자궁암 선별은 가정의, 일반의, 전문 간호사(nurse practitioners)가 하기 때문에 이들을 목표 사용자로 고려하였다.
- **Outcomes** : 일차의료인이 제시된 선별 주기와 선별검사 방법을 얼마나 잘 따르는 지가 주된 결과이다. 의사에 따라 선별검사의 처방 빈도의 변이가 심하다.
- **Healthcare settings and context** : 일차의료

2. 개발 방법의 결정: 신규 개발(*de novo*) vs 수용개작(adaptation)

진료지침 개발을 위해서는 짧게는 몇 개월에서 길게는 몇 년까지 걸리는 과정을 거쳐야 하기 때문에 비용과 인력이 엄청나게 소요되고, 이미 여러 나라에서 다양한 진료지침을 개발하였기 때문에 국제적으로 통용되는 양질의 진료지침이 있는 경우 적절한 방법론으로 해당 진료지침을 수용개작할 수 있다. 진료지침을 신규 개발할 지, 수용개작할 지를 결정하는 기준은 아직 없지만, 아래의 기준을 대부분 만족할 때 수용개작을 고려할 수 있다.

- 정해진 주제와 범위에 부합하는 진료지침이 존재함
- 진료지침의 질이 낮지 않은 경우
- 외국에서 개발된 진료지침을 국내에 수용하는데 문제가 없는 경우
- 기존 진료지침의 최신성에 문제가 없는 경우
- 수용개작 대상인 진료지침 개발자가 수용개작을 반대하지 않는 경우
- 개발 주체가 속한 기구에서 수용개작을 인정한 경우
- 수용개작 방법론에 대한 신뢰가 있는 경우

수용개작하기로 결정하였으면 임상진료지침 수용개작 매뉴얼을 참조한다.

1단계.

진료지침
주제선정

2단계.

진료지침
개발그룹
구성

3단계.

기존
진료지침
검토

4단계.

개발
계획수립

5단계.

핵심질문
결정

3. 진료지침 개발 계획서의 작성

범위와 핵심질문을 포함한 계획서를 작성한다. 가능하면 시간표까지 명시하도록 한다. 다음과 같은 내용으로 한다.

- 서론
- 범위와 핵심질문
- 개발위원 명단, 소속, 해당 전문 분야, 이해상충 선언
- 개발 계획의 업무 계획표(업무, 수행 주체, 책임자, 예상 기간 등)

확인사항	예	아니오
1. 진료지침의 목적이 기술되어 있다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 진료지침의 적용 대상(환자 혹은 인구집단)이 기술되어 있다.	<input type="checkbox"/>	<input type="checkbox"/>
3. 진료지침의 목표 사용자가 기술되어 있다.	<input type="checkbox"/>	<input type="checkbox"/>

제 1부. 기획

5단계. 핵심질문 결정

요약

- 핵심질문을 PICO (Patient/Problem, Intervention, Comparison, Outcomes)의 4가지 요소를 포함하여 구체적인 질문으로 만든다.
- 핵심질문의 목록을 진료지침에 제시한다.

핵심질문 결정은 정해진 범위의 내용을 질문 형태로 구체적이고 명료화시키는 것이다. 이를 바탕으로 근거의 수집 및 검토가 이루어지고 진료지침 권고안이 만들어진다. 적절하고 구체적인 질문을 정의하는 것이 중요하며 보통 기존 진료지침을 검토하고 관련 전문가의 의견을 종합하여 설정하게 된다. 하나의 진료지침에서 다루는 핵심질문의 수는 개발기간, 인력, 비용 등을 고려하여 결정한다. SIGN의 경우 가이드라인을 개발하는 28개월 동안 질문의 수를 40개 이하로 정하며, NICE의 경우 10~18개월 동안 15~20개 핵심질문을 해결하도록 하고 있다.

핵심질문은 PICO (Patient/Problem, Intervention, Comparison, Outcomes)의 4가지 요소를 포함하는 구체적인 형태의 질문으로 만들어야 한다.

- **P**: 질병명, 질병의 단계 및 중증도, 성, 연령 등 특성을 구체적으로 기술한다.
- **I**: 치료법, 진단법, 폭로내용을 정의한다. 치료(폭로)의 종류를 구체적으로 나타내야 하고 약물치료의 경우 가능하면 용량, 요법, 투여시기, 투여경로 등을 명시한다.
- **C**: 중재내용과 비교하는 대조군이 있는지, 있다면 위약인지 다른 중재인지 기술한다.
- **O**: 치료결과가 무엇인지(중기 또는 단기 효과측정, 사망률, 유병률, 합병증, 재발률, 재 입원율, 직장으로의 복귀, 신체적 사회적 기능, 삶의 질, 전반적인 건강수준, 비용)를 기술한다.

예시) PICO를 이용한 핵심질문

과민성 장증후군(IBS) 환자에서
P (patient)

항무스카린제는 위약 또는 치료를 하지 않는 것에 비해
I (intervention) C (comparison)

IBS 증상 장기간 조절에 효과적인가?
O (outcomes)

개발 초기단계에서 결정된 핵심질문이 최종 질문으로 결정되는 경우는 드물고, 일차적인 근거 검색이 이루어진 후 수정되거나 새로운 질문이 생성되어 최종 질문 목록이 확정되는 경우가 흔하다. 만들어진 질문은 진료지침 이해당사자 등의 의견수렴 절차를 거치는 것이 좋다.

확인사항	예	아니오
1. 진료지침이 다루는 핵심질문이 기술되어 있다.	□	□

1단계.
진료지침
주제선정

2단계.
진료지침
개발그룹
구성

3단계.
기존
진료지침
검토

4단계.
개발
계획수립

5단계.
핵심질문
결정

제 2부. 개발

6단계. 근거의 검색

요약

- 검색의 민감도를 최대한으로 높인 포괄적인 검색을 실시한다.
- 검색의 범위에 적절한 데이터베이스를 선택한다.
- 진료지침의 범위에 적절한 검색어로 검색전략을 수립한다.
- 체계적 문헌고찰, 무작위 대조 연구 등의 순서로 문헌검색을 진행한다.

핵심질문에 기초하여 검색전략을 짜는 것이 필요하다. 포괄적인 문헌검색이 되도록 검색할 데이터베이스를 선정하고 데이터베이스별 검색전략과 적절한 검색어의 조합을 만든다. 임상 전문가 및 정보 전문가(information specialist)와 함께 검색어, 검색전략 및 데이터베이스를 결정한다. 실제 검색을 할 때에는 사전 검색(pilot search)을 통해 검색전략을 평가하여 수정할 수 있다.

1. 데이터베이스 선정

일반적으로 문헌검색을 위해 검색원을 선정할 때 미국립의학도서관(National Library of Medicine, NLM)이 제시한 COSI (COre, Standard, Ideal) 모델을 사용한다. 이 중 Core는 문헌 검색의 핵이 되는 부분으로 관련 국내 문헌, 핵심 데이터베이스(예를 들어 PubMed, Embase, Cochrane CENTRAL 등) 등이다. Standard는 표준 검색 범위라고 할 수 있는데 핵심 학술지에 대한 수기 검색, Core에 있는 데이터베이스 이외의 다른 일반적 데이터베이스(Web of Science, DARE 등) 그리고 주제가 특화된 데이터베이스(CINAHL, PsycINFO, ERIC, CANCERLIT 등)를 말하며, Ideal은 이상적인 검색 범위로 학술대회 초록집, 출판되지 않은 문헌, 현재 진행되고 있는 임상시험 등이 포함된다. 대체로 진료지침의 경우 Core를 중심으로 검색이 이루어진다.

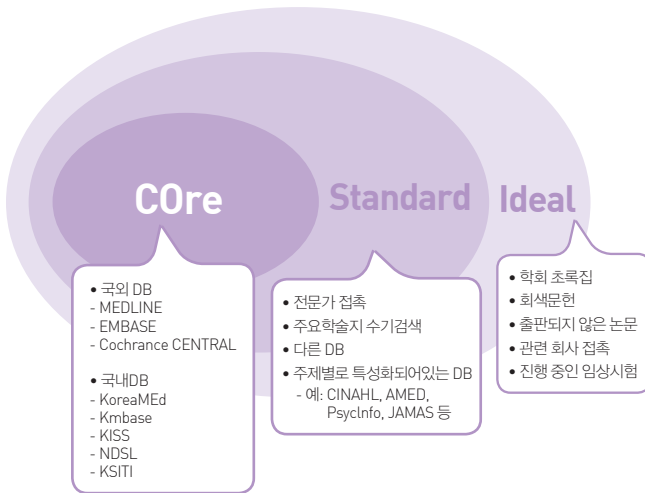


그림 4. COSI 모델에서 제안하는 검색원 범위

권장하는 COre 검색 데이터베이스는 다음과 같다.

a. 국내

DB명	제공기관	URL
■ KoreaMed	대한의학학술지편집인협회	http://www.koreamed.org
■ 한국의학논문데이터베이스	의학연구정보센터	http://kmbase.medic.or.kr
■ 한국학술정보	한국학술정보	http://kiss.kstudy.com
■ 과학기술정보통합서비스	한국과학기술정보연구원	http://www.ndsl.kr
■ 과학기술학회마을	한국과학기술정보연구원	http://society.kisti.re.kr

b. 국외

제공기관	URL	데이터베이스에서 제공하는 자료의 특성
■ Cochrane CENTRAL	http://www.thecochranelibrary.com	controlled trial
■ MEDLINE	http://gateway.ovid.com/autologin	1, 2차문헌
■ EMBASE		Ovid-Interface

주) 한국보건 의료연구원은 Medline, Embase를 Ovid 인터페이스 사용

6단계.
근거의
검색

7단계.
근거의
평가

8단계.
근거의
종합

9단계.
권고안
작성,
권고등급
결정

10단계.
합의안
도출

연구에서 필요하다고 판단되는 경우 다음의 데이터베이스에서 검색하는 것을 고려할 수 있다. 이 외에도 연구 주제에 따라 검색해 볼 수 있는 데이터베이스는 다양할 수 있다.

제공기관	URL	비고
국내		
<ul style="list-style-type: none"> ■ OASIS ■ 한국전통저서포탈 	<p>http://oasis.kiom.re.kr</p> <p>http://www.koreantk.com</p>	한의학문헌
국외		
<ul style="list-style-type: none"> ■ CINAHL 	http://search.ebscohost.com/login.asp?profile=ehost&defaultdb=rzh	간호, 보건의학문헌, 대체의학문헌 DB
<ul style="list-style-type: none"> ■ AMED 	http://gateway.ovid.com/autologin	대체의학문헌 DB
<ul style="list-style-type: none"> ■ PsycINFO 		정신과학문헌 DB
<ul style="list-style-type: none"> ■ JAMAS 	http://www.jamas.or.jp	일본DB, 필요시 (1983-)
<ul style="list-style-type: none"> ■ Medical Online 	http://www.meteo-intergate.com	일본DB, 필요시 [-1983]
<ul style="list-style-type: none"> ■ China Medical Journal 	http://www.cnki.co.kr	중국DB, 필요시

국내 데이터베이스의 경우 검색식을 통한 검색이 불가능하거나 불리언연산자를 인식하지 못하는 검색엔진이 있으므로 검색 후에도 수기검색을 고려해야 한다.

2. 검색어

검색어는 PICO (patient, intervention, comparison, outcomes)에서 도출되며, 주제어(subject headings, MeSH, Emtree 등)와 텍스트 검색을 모두 사용하는 것이 일반적이다.

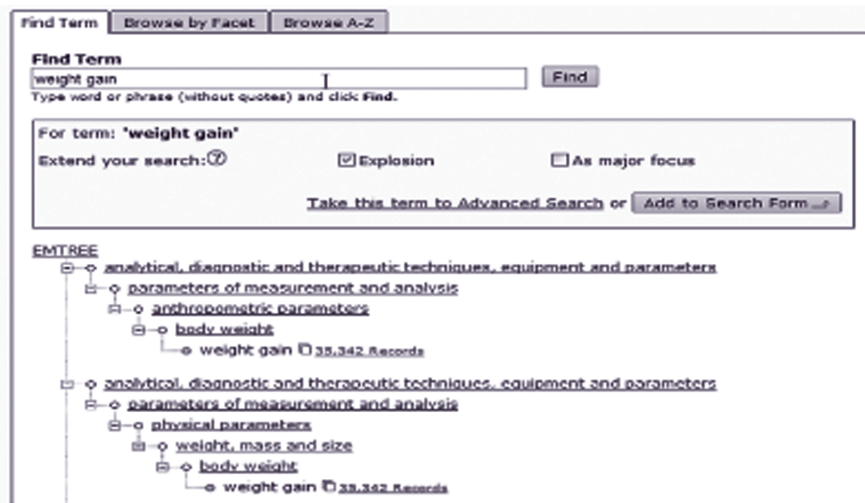


그림 5. EMBASE의 Emtree 예시

주제어는 표준화된 단어나 구의 형태로 문헌의 주요 내용을 대표하는 것으로 tree 형태로 가치를 치고 있어 상·하부에 어떤 주제어가 있는지를 살펴보고 정해야 한다(그림 5).

연구설계에 따른 검색을 실시할 경우 무작위배정 임상시험, 체계적 문헌고찰은 SIGN, PubMed, Cochrane highly sensitive search strategy, InterTASC Information Specialists' Sub-Group, BMJ의 Clinical Evidence 등에서 제시하고 있는 검색전략을 사용해 볼 수 있다. 이밖에도 연구설계에 대한 검색전략을 보고하고 있는 개별 논문을 참조해 볼 수도 있으며 Center for Review and Dissemination (CRD, <http://www.york.ac.uk/inst/crd/intertasc/index.htm>) 홈페이지에서 다양한 검색전략을 모아 수록해 놓고 있으므로 이를 참조하면 된다.

언어편향을 줄이기 위하여 언어제한을 하지 않는 것이 원칙이나 비용, 인력 등을 고려하여 일부 언어만 포함하기도 한다.

3. 검색전략

문헌검색은 연구설계별로 접근하는 것이 좋다. 1단계에서는 체계적 문헌고찰 등 이차문헌을 중심으로 검색하여 검토하며, 해당 임상질문에 대해 충분한 지 판단한다. 만일 그렇지 않다면 무작위 대조 연구, 관찰 연구 순으로 검색을 확대한다(그림 6).

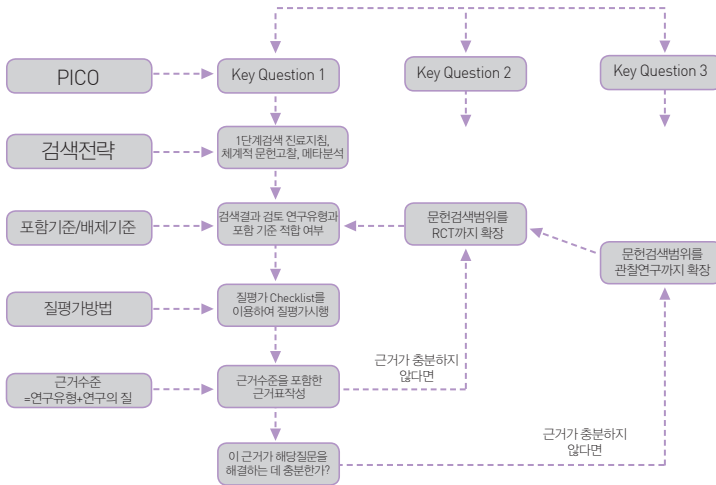
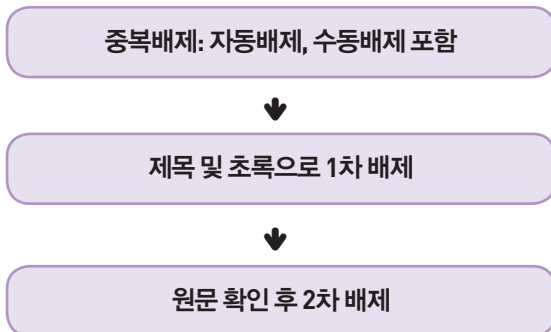


그림 6. 검색과정(SIGN)

4. 문헌 포함/배제

본격적인 문헌 포함 및 배제 작업을 시작하기 전에 **pilot test**를 통해서 추후 선택/배제 과정에서의 시행착오를 줄이는 것을 추천한다.

문헌선정은 적어도 두 명의 연구자가 독립적으로 수행하며, 일반적으로 중복배제, 제목 및 초록 확인 후 배제, 원문 확인 후 배제로 진행된다(그림 7). 이 때 흐름도를 제시하여 독자가 쉽게 이해할 수 있도록 한다(그림 8). 문헌선정 결과에 연구자 간에 불일치가 있는 경우 연구자 간의 토론 혹은 제 3자의 개입을 통하여 불일치를 해소하도록 한다.



◀ 그림 7. 문헌선택 과정

문헌을 배제하는 경우 1차 배제 시에는 주로 해당 사유 및 개수를 흐름도에 간단히 남기고, 2차 배제 시에는 사유를 별도로 정리하고 보고서 작성 시 부록에 보고한다.

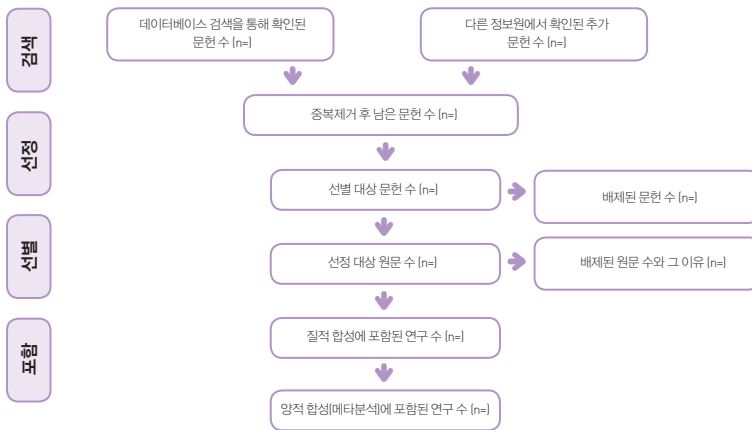


그림 8. 문헌 선택 흐름도(PRISMA flow chart)

검색을 한 이후로 더 추가된 문헌이 있는지를 확인하기 위해 연구 종료시점 1~2개월 전에 재검색을 시행해야 하며, 재검색을 위해 각 검색엔진과 웹사이트에 자동 알림시스템(alerting system)을 이용할 수 있다.

확인사항	예	아니오
1. 문헌 검색을 체계적으로 수행하였다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 근거를 선택하는 기준이 명시되어 있다.	<input type="checkbox"/>	<input type="checkbox"/>

제 2부. 개발

7단계. 근거의 평가

요약

- 근거 평가는 연구설계 분류, 연구설계에 적합한 평가 도구 결정, 근거 평가 순으로 진행된다.
- 연구설계 분류는 study Design Algorithm for Medical literature of Intervention (DAMI)을 이용해서 시행한다.
- 연구설계에 적합한 질 평가 도구를 이용해서 문헌의 질을 평가한다.
- 체계적 문헌고찰은 Assessment of Multiple SysTemAtic Reviews (AMSTAR), 무작위 배정 비교임상시험은 코크란의 Risk Of Bias (ROB) 도구, 비무작위연구는 Risk of Bias Assessment tool for Non-randomised Study (RoBANS)를 이용하여 평가한다.

선택된 문헌에 대해 타당성 평가 결과는 문헌의 근거수준을 결정하므로 권고등급에 영향을 미치게 된다. 근거 평가는 주로 연구설계에 초점을 두며, 연구 결과의 타당성을 손상시키는 비뮴림이 어느 정도인가에 의해 결정된다. 따라서 근거 평가는 연구설계 분류, 연구설계에 적합한 평가 도구 결정, 근거 평가 순으로 진행된다.

1. 연구설계 분류

연구설계에 따라 문헌의 비뮴림 위험 평가를 하기 위해 문헌분류 작업을 수행한다. 이 때 사용될 수 있는 문헌분류도구는 study Design Algorithm for Medical literature of Intervention (DAMI)이다(그림 9). 실제 분류작업을 진행할 때는 이용자 매뉴얼을 이용한다(표 7).

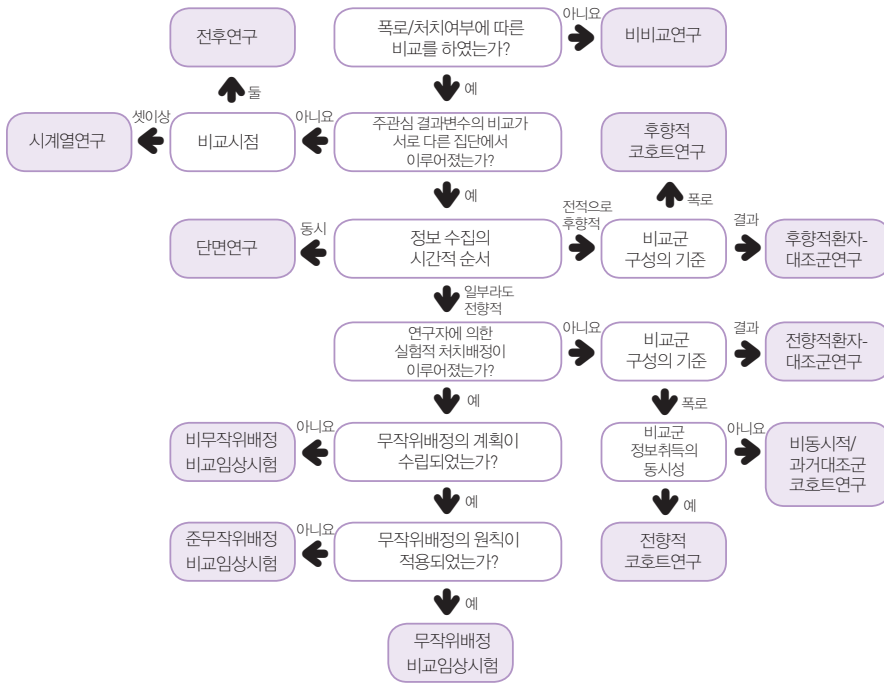


그림 9. 문헌의 분류도구, DAMI

표 7. 문헌의 분류도구, DAMI

• 문헌의 분류도구(DAMI) 이용자 매뉴얼

1) 폭로 또는 중재 여부에 따른 비교를 하였는가?

연구의 주목적이 독립적인 군 사이의 비교(between group comparison, 군간 비교) 또는 동일 군 내에서 둘 이상 시점 사이의 비교(within group comparison, 군내 비교)인 경우 비교연구라고 한다. 반면, 주요변수를 대상으로 군간 비교결과를 제시하더라도 비교연구라는 근거를 뚜렷이 제시하지 못하는 연구는 비교연구로 보지 않는다. 예를 들어 전체 대상자를 성에 따라 구분하여 각각에 대하여 평균 나이를 제시하는 것과 같이 단순히 특정 변수에 대한 기술통계만 제시할 뿐, 통계적 추론이 포함된 정량적 통계분석을 제시하지 않는 연구는 비비교연구의 일종인 환자군 연구(case-series)로 판정한다.

2) 주관심 결과변수의 비교가 서로 다른 집단 사이에서 이루어졌는가?

연구결과에서 제시되는 여러 결과변수에 대해 군간 비교, 군내 비교 모두를 제시하는 경우도 있다. 이 경우 주관심 결과변수를 기준으로 이것이 군간 비교, 군내 비교 중 어디에 해당하는지를 판단한다. 만일 '예'인 경우(군간 비교), 평행연구로 구분되고 '아니오'(군내 비교)인 경우 전후연구로 구분된다. 일부 연구에서 기본적인 연구의 틀은 군내 비교이나, 별도의 독립적인 비교군을 제시하는 경우가 있다. 이 경우 전후연구의 일종인 비교군 전후연구(controlled before and after study)로 구분한다. 다만 교차시험(cross-over trial)인 경우 본 단계를 생략하고 '연구자에 의한 실험적 처치 배정이 이루어졌는가?'라는 질문으로 넘어간다.

3) 정보수집의 시간적 순서

연구에 사용된 중재와 (중재)결과 자료의 수집이 동시에 이루어졌다면 단면연구로 구분한다. 자료가 시간의 흐름에 따라 연속적으로 생성되거나 연구 자료가 전적으로 이미 모두 존재하는 것이라면 후향적 연구로 구분한다. 시간의 흐름에 따라 생성되는 자료이며 아직 발생하기 전 상태인 자료이면 후향적 자료가 일부 포함되어 있더라도 전향적 연구로 본다. 과대평가를 막기 위해 '전향적'인지 판단할 수 없는 경우는 후향적 연구로 본다.

4) 연구자에 의한 실험적 배정이 이루어졌는가?

특정 중재 시행 여부가 연구자의 의도에 따라 행해졌는가를 평가한다. 만일 '예'인 경우 중재연구로 구분하며, 그렇지 않은 경우 관찰연구로 구분한다. 배정 여부가 불확실한 경우 해당 대상자에게 연구 당시의 지식에 근거해 통상적으로 행해지는 임상적인 처치(예: 폐렴에서 특정 항생제 요법)는 실험적 처치라고 보지 않으며, 특정 처치의 유효성 또는 안전성에 대한 결론을 내리지 못한 상태에서 이를 적용하는 경우에만 실험적 처치로 판단한다. 특정 처치의 실험적인 성격이 명백하지 않은 경우에는 평가자의 합의로 결정한다.

5) 무작위로 배정되었다는 언급이 있는가?

연구방법에 무작위배정에 대한 언급이 있는지 살핀다. 환자나 연구자의 선호도에 의해 배정되었다면 '아니오'에 해당된다. 무작위배정에 대한 언급이 불확실한 경우에도 '아니오'로 판단한다.

6) 무작위배정의 원칙이 적용되었는가?

배정방법이 진정한 무작위배정의 원칙에 부합하는 것인지를 판단한다. 예를 들어 생일, 병록 번호, 내원순서 등에 따른 중재군의 배정은 진정한 무작위배정의 원칙에 부합하지 않기 때문에 '아니오'로 판단하고 준무작위배정 비교임상시험으로 구분한다.

2. 문헌의 질 평가

다양한 질 평가도구가 있으나 모든 연구설계에 적용할 수 있는 평가도구는 현재 존재하지 않는다. 그래서 연구설계별 질 평가도구를 선정하여 문헌의 질 평가를 하는 것이 바람직하다. 체계적 문헌고찰은 Assessment of Multiple SysTemAtic Reviews (AMSTAR), 무작위배정 비교임상시험은 코크란의 Risk Of Bias (ROB) 도구, 비무작위연구는 Risk of Bias Assessment tool for Non-randomised Study (RoBANS)를 이용하여 평가한다.

1) 체계적 문헌고찰

체계적 문헌고찰의 질 평가는 AMSTAR (부록 2)를 사용한다. 체계적 문헌고찰의 질 평가는 잘 수행된 체계적 문헌고찰을 선택하는 중요한 과정이다.

질 평가는 최소한 2명 이상이 수행하여야 하며 의견이 다를 때 이를 해소할 수 있는 방법을 명시한다.

AMSTAR는 총 11개 항목을 평가하도록 이루어진 체크리스트 방식의 도구이며, 타당성 검증(validation)이 된 도구이다. 평가자는 각 항목에 대하여 '예', '아니오', '대답할 수 없음', '적용할 수 없음' 네 가지로 대답한다. '예'는 해당항목에 관련하여 체계적 문헌고찰이 잘 다루었다는 것을 의미하며, '아니오'는 잘 수행되지 않았음을 의미한다. '대답할 수 없음'은 해당항목을 시행했는지 알 수 없는 경우를 의미하며, '적용할 수 없음'은 해당항목을 적용할 수 없어 수행하지 못한 경우라 할 수 있다.

AMSTAR 결과를 이용하는 방법은 다양하고, 합의된 방법은 없으므로 연구진에서 결정할 필요가 있다. AMSTAR를 점수화하여 0-3점이면 낮은 질, 4-7점이면 중등도의 질, 8-11점은 높은 질로 분류하는 방법을 추천한다.

AMSTAR로 체계적 문헌고찰의 질을 평가하여 일정 정도의 질을 가진 것으로 판단될 때는 해당 체계적 문헌고찰은 진료지침 개발을 위한 근거로 사용된다. 일정 정도가 어느 정도인지는 현재 합의된 내용이 없으므로 진료지침 개발그룹에서 합의하여 결정한다. 만일 일정 정도 이상의 질을 가진 체계적 문헌고찰이 없는 경우에는 새로운 체계적 문헌고찰을 시행할 수 있다.

2) 무작위배정 비교임상시험

무작위배정 비교임상시험에서의 비뚤림 종류는 선택 비뚤림, 수행 비뚤림, 탈락 비뚤림, 결과 확인 비뚤림, 보고 비뚤림의 다섯 가지이다. 여러 무작위배정 비교임상시험 평가도구들이 있으나 Cochrane group이 개발한 ROB 도구를 많이 사용한다(부록 3).

ROB 도구는 연구의 비뚤림 위험을 평가하는데 초점을 두고 있다. 이 도구는 체크리스트 방식과 영역평가 방식이 결합된 형태이며, 문항을 단순화한 반면에 같은 문항에 대한 주관적이거나 자의적인 답변의 가능성을 최소화하였다. 그래서 평가자의 연구방법론에 대한 이해도와 숙련도에 따른 평가결과의 변동을 최소화할 수 있다. 각 문항은 비뚤림 위험이 '높음', '낮음', '불확실'로 판정하고, 구체적인 가이드라인이 제시되어 있어 사용이 용이하며, 근거수준을 평가할 때 GRADE에 적용하기 편리한 장점이 있다.

① 평가영역

ROB 도구는 무작위배정순서가 적절한 방법에 의해 생성되었는지, 배정순서가 적절히 은폐되었는지, 참여자·연구자·결과평가자에 대한 눈가림이 적절한지, 불완전한 자료에 대한 분석이 적절한지, 제시된 자료가 선택적 보고일 가능성이 없는지, 다른 비뚤림의 위험은 없는지에 대한 6개의 영역으로 구성되어 있다. 이 6개 영역을 통하여 평가할 수 있는 비뚤림 위험 종류와의 관계는 표 8에서 확인해 볼 수 있다.

표 8. 비뚤림 종류와 관련 영역

비뚤림 종류	기술	Risk of bias에서의 관련 영역
선택 비뚤림 (selection bias)	비교 대상군들의 기저상태의 체계적인 차이 (systematic differences)	<ul style="list-style-type: none"> 무작위 배정순서 생성 배정순서 은폐
실행 비뚤림 (performance bias)	제공되는 치료가 계통적인 차이가 있거나, 관심 중재 이외의 다른 인자에 노출되는 경우	<ul style="list-style-type: none"> 연구 참여자, 연구자에 대한 눈가림 타당도를 위협하는 다른 잠재적 비뚤림

결과 확인 비뚤림 (detection bias)	(중재)결과를 확인하는 방법에서의 그룹간 체계적인 차이	<ul style="list-style-type: none"> • 결과 평가에 대한 눈가림 • 타당도를 위협하는 다른 잠재적 비뚤림
탈락 비뚤림 (attrition bias)	탈락률에 있어서 그룹간 체계적인 차이	<ul style="list-style-type: none"> • 불완전한 결과 자료
보고 비뚤림 (reporting bias)	보고된 결과와 보고되지 않은 결과간의 체계적인 차이	<ul style="list-style-type: none"> • 선택적 결과 보고

② 평가방법

평가자는 각 항목에 대하여 비뚤림 위험이 '높음(high risk of bias)', '낮음(low risk of bias)', '불확실(uncertain risk of bias)' 세 가지로 대답한다. '낮음'은 비뚤림의 해당항목에 대해 비뚤림의 가능성이 낮음을 의미하며, '높음'은 높은 비뚤림 가능성을 의미한다. '불확실'한 경우는 비뚤림에 대한 위험을 판단하기 어려운 경우이다.

각 문항에 대하여 '낮음', '높음', '불확실'로 응답 결과를 RevMan에 입력하여 그림 형태로 나타내면, 체계적 문헌고찰에 포함된 문헌들의 각 평가영역에 대한 전반적인 평가결과를 파악할 수 있다(그림 10, 그림 11).

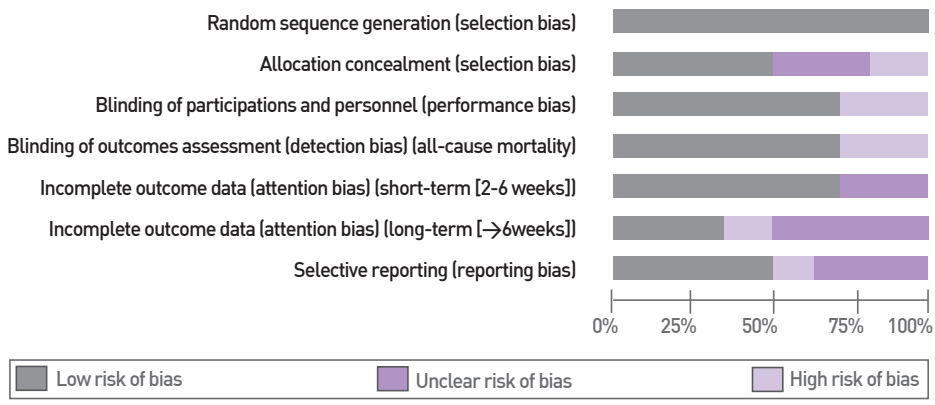


그림 10. 비뚤림 위험 평가결과 예시(전체 결과) (Cochrane handbook, 2011) Reproduced table. Copyright © 2011 The Cochrane Collaboration, Published by John Wiley & Sons, Ltd

③ ROB 도구 해석방법

비뚤림 위험의 해석은 네 가지 방법으로 해석될 수 있다.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (selection bias) (patient-reported outcomes)	Blinding of outcome assessment (selection bias) (all-cause mortality)	Incomplete outcome data (attrition bias) (short-term < 5 weeks)	Incomplete outcome data (attrition bias) (long-term > 6 weeks)	Selective reporting (reporting bias)
Barr 1998	+	+	+	+	+	+	+	+
Baylis 1999	+	+	+	+	+	+	+	+
Cooper 1987	+	?	+	+	+	+	+	+
Dodd 1995	+	?	+	+	+	+	+	+
Goodwin 1996	+	+	+	+	+	+	+	+
Sanders 1993	+	+	+	+	+	+	+	+

그림11. 비뚤림위험 평가결과 예시(연구별) (Cochrane handbook, 2011) Reproduced table. Copyright © 2011 The Cochrane Collaboration, Published by John Wiley & Sons, Ltd

㉠ 한 연구에서 결과 간의 비뚤림 위험 요약

순서생성이나 배정은폐와 같은 영역은 한 연구 안에서 여러 결과에 영향을 주지만, 눈가림이나 불완전 결과 보고는 결과에 따라 다른 비뚤림 위험을 갖게 된다. 따라서 연구 내 모든 결과에 대해 하나의 비뚤림 위험으로 나타내는 것은 일반적으로 추천되지 않는다.

㉡ 한 연구에서 하나의 결과의 비뚤림 위험을 요약

한 연구에서 결과에 따라 비뚤림 위험이 다를 수 있기 때문에 이 방법으로 요약하는 것이 추천될 수 있다. 비뚤림 위험 요약에 위해서는 결과에 따라 모든 항목(예를 들어 연구 단계 평가 항목인 배정은폐와 결과에 따른 평가 항목인 눈가림 등)을 포함해야 한다.

㉢ 연구 간에 하나의 결과에 대해 비뚤림 위험을 요약(예: 메타분석)

메타분석이나 GRADE 평가에서 주로 이용되는 방식이다.

㉔ 모든 연구와 모든 결과에서 비뚤림 위험을 요약

이 요약은 두 가지 이유로 시행되지 말아야 한다. 첫째, 어떤 결과가 의사결정에 중요한 지에 대한 가치 판단이 요구된다. 그러나 부작용과 같은 중요할 수 있는 결과에 대한 자료는 평가에 포함된 몇몇 연구들에서 접근가능하지 않을 수 있다. 그러므로 비뚤림 위험은 모든 결과들에서 동일하지 않다. 둘째, 어떤 결과가 의사결정에 중요한 지에 대한 판단은 가치와 기저 위험도와 같은 다른 인자들로 인해 임상상황에 따라 다를 수 있다. 따라서 체계적 문헌고찰의 맥락에서는 권유하지 않는다. 한 연구에서 비뚤림 위험의 의미와 연구간 비뚤림 위험의 의미 차이는 아래 표 9에서 확인할 수 있다.

표 9. 연구 내 및 연구 간 비뚤림 위험 요약 접근 방법

비뚤림 위험	해석	연구내 (Within a study)	연구간 (Across studies)
낮음	심각할 것 같지 않은 비뚤림 가능성	모든 주요 영역에서의 낮은 비뚤림 위험	연구들에서 나온 대부분의 정보가 낮은 비뚤림 위험
불확실	결과에 대한 의심을 불러일으키는 비뚤림 가능성	하나 또는 그 이상의 주요 영역에서의 비뚤림 위험 불확실함	연구들에서 나온 대부분의 정보가 낮거나 불확실한 비뚤림 위험
높음	결과의 신뢰도를 심각하게 약화시키는 비뚤림 가능성	하나 이상의 주요 영역에서의 비뚤림 위험 높음	높은 비뚤림 위험의 연구에서 얻은 정보 비율이 높아 결과 해석에 영향을 미치지 충분함

2) 비무작위 연구

비무작위 연구에서 비뚤림의 종류는 선택 비뚤림, 수행 비뚤림, 탈락 비뚤림, 결과 확인 비뚤림, 선택적 보고 비뚤림의 다섯 가지이다. 비무작위 연구의 비뚤림 위험을 평가하는 도구는 매우 다양하다. RoBANS (Risk of Bias Assessment tool for Non-randomised Study)는 타당성 검증이 이루어졌고, 비무작위배정 비교임상시험 연구, 코호트 연구, 환자-대조군 연구, 전후 연구의 평가에 쓰인다.

RoBANS는 비무작위 연구의 질 평가를 위해 2009년 개발된 도구이다(부록 4). ROB 도구와 유사하게 체크리스트 방식과 영역평가 방식이 결합된 형태를 갖추고 있다. 비뚤림을 종류에 따라 구분하고 해당 비뚤림 위험을 평가할 수 있는 관련 영역을 설정하였다. 그리고 해당 영역을 측정할 수 있는 항목을 정의하고 판단 기준을 제시하였다. 판단 기준은 각 항목에 대하여 개별 결과의 비뚤림 위험을 어떻게 판정할 지에 대하여 기술하고 있다.

① RoBANS 평가 영역

비무작위 연구에서 5가지 비뚤림을 대상군 선정, 교란변수, 노출 측정, 불완전한 결과 자료, 결과 평가 눈가림, 선택적 결과 보고의 6가지 영역으로 평가한다(표 11).

표 10. 비뚤림의 종류와 평가 영역

비뚤림 종류	무작위배정 비교임상연구	비무작위 연구
선택 비뚤림 (selection bias)	<ul style="list-style-type: none"> • 무작위 배정순서 생성 • 배정순서 은폐 	<ul style="list-style-type: none"> • 대상군 선정 • 교란변수
실행 비뚤림 (performance bias)	<ul style="list-style-type: none"> • 연구 참여자, 연구자에 대한 눈가림 • 타당도를 위협하는 다른 잠재적 비뚤림 	<ul style="list-style-type: none"> • (중재)노출에 대한 측정
결과 확인 비뚤림 (detection bias)	<ul style="list-style-type: none"> • 결과 평가에 대한 눈가림 • 타당도를 위협하는 다른 잠재적 비뚤림 	<ul style="list-style-type: none"> • 결과 평가 눈가림
탈락 비뚤림 (attrition bias)	<ul style="list-style-type: none"> • 불완전한 결과 자료 	<ul style="list-style-type: none"> • 불완전한 결과 자료
보고 비뚤림 (reporting bias)	<ul style="list-style-type: none"> • 선택적 결과 보고 	<ul style="list-style-type: none"> • 선택적 결과 보고

표 11. RoBANS 도구 평가영역

영역	설명	비뚤림 위험	판단근거 (논문에서 그대로 인용함)
대상군 선정	부적절한 대상군 선정으로 인해 발생한 선택 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	
교란변수	교란변수 확인과 고려가 부적절하여 발생한 선택 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	
중재(노출) 측정	부적절한 중재(노출) 측정으로 인해 발생한 실행 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	
결과 평가 눈가림	부적절한 결과 평가 눈가림으로 인해 발생한 결과 확인 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	
불완전한 결과 자료	불완전한 자료를 부적절하게 다루어 발생한 탈락 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	
선택적 결과 보고	선택적 결과 보고 때문에 발생한 보고 비뚤림	<input type="checkbox"/> 낮음 <input type="checkbox"/> 높음 <input type="checkbox"/> 불확실	

② 평가방법

평가자는 도구를 이용할 때 각 부분 항목에 대하여 ‘낮음’, ‘높음’, ‘불확실’ 세 가지로 대답한다. ‘낮음’은 비플립의 해당 항목에 대해 비플립의 가능성이 낮음을 의미하며, ‘높음’은 높은 비플립 가능성을 의미한다. 불확실한 경우는 비플립에 대한 위험을 판단하기 어려운 경우이다. 각 문항에 대하여 낮음, 높음으로 응답된 결과를 통하여 체계적 문헌고찰에 포함된 문헌들의 각 평가영역에 대한 전반적인 평가결과를 파악할 수 있도록 개별 문헌에 대한 평가결과를 RevMan에 입력하면 그림 형태로 나타낼 수 있다.

③ RoBANS 해석방법

RoBANS의 해석 역시 무작위연구의 평가도구인 ROB 도구와 유사하게 네 가지 방식으로 평가하게 된다.

- 한 연구에서 결과간의 비플립 위험 요약
- 한 연구에서 하나의 결과의 비플립 위험을 요약
- 연구 간에 하나의 결과에 대해 비플립 위험을 요약 (예: 메타분석)
- 모든 연구와 모든 결과에서 비플립 위험을 요약

확인사항	예	아니오
1. 연구설계에 따라서 문헌의 질을 평가하였다.	□	□

제 2부. 개발

8단계. 근거의 종합

요약

- 근거의 종합을 위해 개별 연구에 대해서 연구 특성, 연구 대상자, 중재, 결과, 문헌의 질 평가 등을 내용으로 하는 근거표를 작성한다.
- 주요 결과 각각에 따라 전반적 근거를 요약한 근거 요약표를 작성할 수 있다.

1. 근거표 작성

근거표는 모든 개별 연구들의 내용과 질을 종합하여 하나의 근거로 평가하도록 정리한 것이다. 근거표는 연구에 따라 매우 다양할 수 있다. 일반적으로 연구의 일반정보, 연구 특성, 대상자 특성, 중재, 비교중재, 중재결과(outcomes), 효과측정치, 문헌의 질 평가 결과, 평가자 코멘트 등을 기록한다(표 12).

표 12. 근거표에 포함되는 항목

일반정보	• 자료 추출을 수행한 연구자명/자료 추출 시행 날짜 • 연구 관련 : 번호, 저자, 제목, 서지사항
연구 특성	• 연구설계
대상자 특성	• 연구 시작 시점의 대상자 특성 • 각 특성별 대상자 수
중재	• 중재군(용량, 투여경로, 시행횟수 등)
비교중재	• 대조군 기술(용량, 투여경로, 시행횟수 등)
중재결과	• 결과변수의 종류, 측정 도구, 측정 단위 등
효과측정치	• 중재군, 대조군 : 참여 대상자 수, 분석에 포함된 대상자 수, 결과요약 등 • 연구분석 결과 : OR, RR 등

작성자는 2명이 독립적으로 진행하는 것이 원칙이나, 1인이 작성하고 다른 1인이 확인할 수 있다. 근거표 양식을 만들 때는 가급적 **pilot test**를 진행하여 최종 양식을 확정하도록 한다. 임상연구에서 최종 결과인 생존율이나 삶의 질 등 자료가 없어 대리지표(surrogate marker)로 대신하기도 하는데, 이 지표가 최종 결과와 얼마나 긴밀한 연관이 있는 지를 고려해야 한다.

2. 근거요약표

근거 요약표(summary of findings table, SoF)는 체계적 문헌고찰을 통해 얻은 주요 결과 각각에 대해 전반적인 근거를 요약하여 제시한 표이다. 근거 요약표는 결과 제시표와는 다르며, 근거 요약표에는 근거수준, 중재의 효과 크기, 주요 결과에 대한 자료 요약 등이 제시된다. Cochrane Center For Systematic Review에서는 근거 요약표를 서론 앞에 제시하며, 결과와 고찰 사이에 제시하는 단체도 있다. 요약표는 6가지 항목으로 구성된다.

- 주요 (중재)결과지표: 이득 관련 결과지표, 위해 관련 결과지표
- 군별 위험도: 중재군과 대조군의 사건 발생 위험
- 효과추정치: 절대효과 추정치, 상대효과 추정치
- 연구 수와 참여자 수
- 근거수준
- 비고(comments)

근거 요약표는 직접 만들 수도 있으며 RevMan, GRADE profiler 프로그램을 활용하여 작성할 수 있다. RevMan에 자료를 입력한 경우라면 GRADE로 직접 편입하여 요약표를 만들 수도 있으며, 반대로 GRADE에서 RevMan으로 자료를 편입하여 만들 수도 있다. 아래에 각각의 예시를 제시하였다(그림 12, 그림 13).

Self management for patients with chronic obstructive pulmonary disease						
Patient or population: patients with chronic obstructive pulmonary disease						
Settings: primary care, community, outpatient						
Intervention: self management						
Comparison: usual care						
Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk usual care	Corresponding risk self management				
Quality of Life St George's Respiratory Questionnaire. Scale from: 0 to 100 (follow-up: 3 to 12 months)	The mean quality of life ranged across control groups from 38 to 60 points	The mean Quality of Life in the intervention groups was 2.58 lower (5.14 to 0.02 lower)		698 (7)	⊕⊕⊕○ moderate	Lower score indicates better quality of life A change of less than 4 points is not shown to be important to patients.
Dyspnea Borg Scale. Scale from: 0 to 10 (follow-up: 3 to 6 months)	The mean dyspnoea ranged across control groups from 1.2 to 4.1 points	The mean Dyspnoea in the intervention groups was 0.53 lower (0.96 to 0.1 lower)		144 (2)	⊕⊕○○ low	Lower score indicates improvement
Number and severity of exacerbations	See comment	See comment	Not estimable	591 (3)		Effect is uncertain
Respiratory-related			OR 0.64	966	⊕⊕⊕○	

그림 12. GRADEpro를 이용하여 만든 근거 요약표

Self management for patients with choronic obstructive pulmonary disease

Patient or population: patients with chronic obstructive pulmonary disease

Settings: primary care, community, outpatient

Intervention: self management

Comparison: usual care

Outcomes	Illustrative comparative risks* [95% CI]		Relative effect [95% CI]	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	usual care	self management				
Quality of Life St George's Respiratory Questionnaire. Scale from: 0 to 100 (follow up: 3 to 12 months)	The mean quality of life ranged across control groups from 38 to 60 points	The mean Quality of Life in the intervention group was 2.58 lower [5.14 to 0.02 lower]		698 (7)	⊕⊕⊕○ moderate	Lower score indicates better quality of life A change of less than 4 points is not shown to be important to patients.

그림 13. GRADEpro에서 만든 근거 요약표를 RevMan으로 편집한 결과

제 2부. 개발

9단계. 권고안 작성, 권고등급 결정

요약

- 근거수준은 현재까지의 근거로 해당 중재의 효과에 대해 확신하는 정도이다. 근거수준 평가는 GRADE에 준해서 시행하되 분야, 목적 등에 맞게 수정되어 사용될 수 있다.
- 근거수준은 ① 비뚤림 위험, ② 이질성(일관성 없음), ③ 비직접성, ④ 비정밀, ⑤ 출판 비뚤림이 있는 경우 근거수준 정도를 1단계 혹은 2단계를 낮추고, 만일 ① 효과의 크기가 큼, ② 교란변수의 영향, ③ 양-반응 관계가 있으면 단계를 높인다. 최종적으로 '높음', '중등도', '낮음', '불충분'의 네 단계로 근거수준을 평가한다.
- 권고등급은 권고 대상 환자에서 중재를 시행하였을 때 원하는 결과가 원하지 않는 결과보다 많이 발생할 것으로(혹은 적게 발생할 것으로) 확신하는 정도를 말한다. 권고등급은 일반적으로 ① 근거수준, ② 이득과 유해의 저울질, ③ 가치와 선호도, ④ 자원 이용(비용)의 네 가지를 고려하여 결정한다.
- 권고등급은 strong for: A, weak for: B, weak against: C, strong against: D, insufficient: I로 분류한다.
- 권고안은 특이적이며 모호하지 않아야 한다. 주요 권고안은 쉽게 확인할 수 있어야 하며, 권고와 이를 지지하는 근거들이 명시적으로 연결되어 있어야 한다. 또한 권고등급이 제시되어야 한다.

어떤 핵심질문에 대한 근거를 파악하고 평가한 후에는 전체 근거를 종합해서 권고안을 도출하고 권고등급을 결정한다. 따라서 권고안 작성은 다음의 순서로 진행된다.

첫째, 주요 결과에 대해서 근거수준을 결정한다.

둘째, 해당 중재의 주요결과를 전체적으로 판단하여 근거수준을 결정한다.

셋째, 근거수준을 바탕으로 이득과 유해의 저울질, 가치, 비용 등을 판단하여 권고등급을 결정한다.

넷째, 권고등급에 기반을 두어 권고안을 작성하고 보고한다.

6단계.
근거의
검색

7단계.
근거의
평가

8단계.
근거의
종합

9단계.
권고안
작성,
권고등급
결정

10단계.
합의안
도출

1. 근거수준과 권고등급 평가

1) 근거수준 평가

근거수준(level of evidence)은 현재까지 근거로 해당 중재의 효과에 대해 확신하는 정도를 말한다. 근거수준은 여러 가지 용어로 쓰이지만 국내에서는 가장 널리 쓰이는 ‘근거수준’이라는 용어를 사용한다. 근거수준을 결정할 때 고려되는 요소로는 ① 연구설계, ② 문헌의 질, ③ 근거의 양, ④ 근거의 일관성, ⑤ 근거의 직접성 등이 있다.

다양한 기구에서 사용하고 있는 근거수준 체계는 근거수준의 요소를 일부 혹은 전체를 사용하여 평가한다. 대체로 개별 학회, 과거에 개발된 체계의 경우 연구설계와 문헌의 질 등 간단한 요소만으로 근거수준을 결정한다. 반면 국가 기구, 최근에 체계를 새로 개발한 경우는 요소 전체를 모두 적용한다. 널리 쓰이는 체계 중 SIGN을 제외하고 GRADE, AHRQ, USPSTF는 모두 근거수준의 모든 요소를 사용한다.

2) 권고등급 평가

권고등급은 권고 대상 환자에서 중재를 시행하였을 때 원하는 결과(desirable effects)가 원하지 않는 결과(undesirable effects)보다 많은 것으로(혹은 적은 것으로) 확신하는 정도를 말한다. 권고등급은 일반적으로 ① 근거수준, ② 이득과 위해의 저울질, ③ 가치와 선호도, ④ 자원 이용(비용)의 네 가지를 고려하여 결정한다.

근거수준과 권고등급의 평가는 GRADE 과정을 사용하여 평가한다. GRADE를 등급화 체계로 사용하는 이유는 아래와 같다.

- 근거수준과 권고강도를 명백히 구분한다.
- 근거수준 결정의 객관적인 기준을 사용한다.
- 근거수준 평가 과정을 모두 공개할 수 있어서 투명화에 기여한다.
- 50개 이상 국제기구에서 적용하고 있다.
- 다양한 적용 방법과 교육 자료 등 풍부한 자료가 존재한다.
- GRADEpro를 이용하여 근거수준 평가 과정을 객관적으로 진행할 수 있다.

GRADE는 체계적인 근거수준 평가 방법론과 권고강도 결정법을 제시한다. 하지만 모든 기관이 GRADE에서 제시한 방법론대로 평가하지는 않고 일부 기관이나 국가는 상황이나 수준에 맞도록 내용을 수정하여 사용한다. 예를 들어 NICE (영국), CADTH (캐나다)의 경우 Evidence profile 만을 이용하고 AHRQ (미국), WHO의 경우 과정을 명확화, 간결화하여 사용하고 있다.

3) GRADE

The GRADE 방법은 Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group (www.gradeworkinggroup.org)에 의해 개발된 것이다. 방법론학자, 진료지침 개발자 등이 근거의 수준과 권고강도의 등급화를 위해 제시한 것이다. 근거에 기반하여 권고안 도출에 이르는 GRADE과정은 아래와 같다(그림 14).

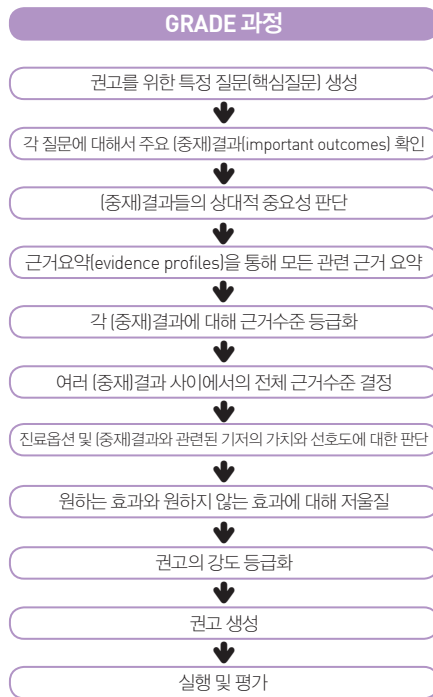


그림 14. GRADE 과정

GRADE의 주요 내용은 아래와 같다.

- 결과에 따른 등급화 : GRADE는 개별 결과에 대해서 근거수준을 결정한다.
- 결과의 중요도 결정 : 결과의 중요도에 따라서 근거수준을 결정한다. 결과는 의사 결정에 핵심적, 결정적(critical) 결과, 중요한(important) 결과, 중요하지 않은(not important) 결과로 구분한다. 근거 요약은 중요한 결과에 대해서만 하며 권고는 핵심적인 결과만을 고려하여 결정한다.
- GRADEpro를 이용한 Evidence profile 작성: GRADE 과정을 거쳐 작성한 체계적 문헌고찰의 결과는 GRADEpro를 이용해서 Evidence profile (그림 15)로 작성한다.

Quality assessment							Summary of findings				Quality	Importance
No. of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	other considerations	No of patients		Effect			
							glucosamine	no treatment	Relative (95% CI)	Absolute		
symptomatic deep vein thrombosis (DVT)												
9	randomised trial	serious	serious ¹	no serious indirectness	serious ²	none	555	556	-	MD -0.19 [-0.5 to 0.11] [§]	⊕⊕○○ LOW	CRITICAL

그림 15. Evidence profile

GRADEpro는 <http://www.cc-ims.net/revman/other-resources/gradepr> 에서 제공하며 무료이다. GRADEpro에 대한 이용 방법이나 GRADE approach에 대한 사항은 help 메뉴에 잘 정리되어 있다.

(1) 근거수준의 결정

GRADE에서 근거수준은 연구설계에 따라 우선적으로 결정된다. 무작위배정 비교임상 시험의 경우 '높음(high)', 관찰연구인 경우 '낮음(low)', 환자군 연구인 경우 '매우 낮음(very low)'로 배정한다. GRADE에서 근거수준의 의미는 표13과 같다.

표 13. 근거수준의 의미

등급	정의
높음 (high)	추후 연구로 효과의 추정치에 대한 확신정도가 바뀔 가능성은 매우 낮다.
중등도 (moderate)	추후 연구로 효과추정치에 대한 확신 정도에 중요한 영향을 미칠 것이며 추정치가 바뀔 수도 있다.
낮음 (low)	추후 연구는 효과 추정치의 확신 정도에 중요한 영향을 미칠 가능성이 매우 크며 추정치가 바뀔 가능성이 높다.
매우 낮음 (very low)	효과에 대한 어떤 추정도 불확실하다.

체계적 문헌고찰을 통한 근거의 평가를 통해 1) 비풀림 위험, 2) 이질성(일관성 없음), 3) 비직접성, 4) 비정밀, 5) 출판 비풀림이 있는 경우 근거수준 정도를 1등급 혹은 2등급 낮춘다. 만일 1) 효과의 크기가 큼, 2) 교란변수의 영향, 3) 양-반응 관계에 따라서 근거수준의 정도를 올릴 수 있다.

(2) 권고강도의 결정

권고등급은 일반적으로 ① 근거수준, ② 이득과 위해의 저울질, ③ 가치와 선호도, ④ 자원 이용(비용)의 권고강도 요인을 고려하여 결정한다. 각 권고강도 요인의 의미는 아래와 같다 (표 14). GRADE에서는 ‘강한’, ‘약한’으로 권고등급을 분류한다.

표 14. 권고강도 요인의 의미

권고강도를 높이는 요인	의미
원하는 효과와 원하지 않는 효과 사이의 저울질	원하는 효과와 원하지 않는 효과 사이의 차이가 크면 강한 권고등급일 가능성이 높아진다. 총 이득이 작고 이득에 대한 불확실성이 클수록 약한 권고등급을 받을 가능성이 커진다.
근거수준	근거수준이 높으면 강한 권고등급을 받는다.
가치와 선호도	가치와 선호도의 변이가 크거나 불확실성이 있으면 약한 권고등급을 받게 된다.
자원 이용(비용)	중재의 비용이 커서 자원 소비가 많아지면 강한 권고등급을 받기 어려워진다.

이상의 GRADE 과정을 그림으로 나타내면 그림16과 같다.

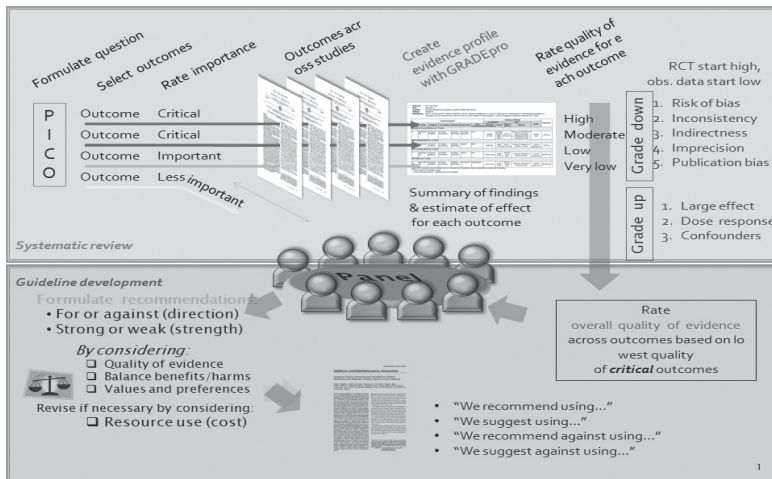


그림 16. GRADE 과정 흐름도

(3) GRADE adaptation

GRADE에서 근거수준은 연구설계, 비뮌립 위험, 이질성(일관성 없음), 비직접성, 비정밀, 출판 비뮌립, 효과의 크기, 교란변수, 양-반응 관계 등으로 결정된다. 이들 요인의 내용 중 세부 지침이 필요하거나 방법론적으로 단순화시킬 필요가 있는 부분에 대하여 기관별로 결정하여 사용할 수 있다. 아래 내용은 한국보건의료연구원에서 사용하기 위한 GRADE 적용기준이다.

① 근거수준

㉑ 연구설계

GRADE	한국보건의료연구원
<ul style="list-style-type: none"> • 무작위배정 비교임상시험: High • 관찰연구: Low • 환자군 연구: Very low 	<ul style="list-style-type: none"> • 높음 : 무작위배정 비교임상시험 • 낮음 : 비무작위배정 비교임상시험, 코호트 연구, 환자-대조군 연구, 단면 연구, 대조군이 있는 전후 연구, 시계열 연구 • 불충분 : 대조군이 없는 전후 연구, 비비교 연구, (전문가 의견)
<ul style="list-style-type: none"> • 설명: GRADE는 연구설계 적용에 불확실성 있어서 체계적 문헌고찰의 대상이 되는 연구설계를 명백히 구분함. 전문가 의견은 GRADE에서는 근거의 범주로 포함하고 있지 않지만, 경우에 따라 포함할 수도 있다. 	

㉠ 비뚤림 위험

GRADE	한국보건의료연구원
<ul style="list-style-type: none"> • 질 평가 영역을 이용하여(특정 도구를 지정하지 않음) 심각한 제한이 있으면 -1, 매우 심한 제한이 있으면 -2를 한다. • 무작위배정 비교임상시험연구는 순서생성, 배정은폐, 눈가림, 불완전 결과, 선택적 보고, 잠재적인 다른 비뚤림 위험을 평가하며, 관찰연구의 경우 대상군선정, 교란변수, 노출과 결과 측정에 대해 평가한다. 	<ul style="list-style-type: none"> • 무작위배정 비교임상시험의 경우 Cochrane ROB를 이용하고, 비무작위 연구의 경우 RoBANS를 이용하여 평가한다. 만일 다른 도구를 이용할 경우 해당 도구를 사용한 이유를 밝힌다. • 결과에 따라 주요 비뚤림 위험 영역을 선정한다(예: 통증 등 주관적 결과 지표인 경우 눈가림을 중요하게 평가함). 연구의 대부분이 비뚤림 위험이 낮다면 낮추지 않고, 자료의 대부분이 불확실하다면 잠재적 제한점이 치료 효과의 추정치를 낮출 수 있는지 여부에 따라서 낮추지 않거나, 한 등급 낮춘다. 대부분의 연구가 비뚤림 위험이 높다면 제한점 정도에 따라서 한 등급 혹은 두 등급 낮춘다. • 전체 연구 중에서 비뚤림 위험이 낮은 연구가 30-50% 이면 한 등급을 낮추고, 30% 보다 적으면 두 등급을 낮춘다.
<ul style="list-style-type: none"> • 설명: GRADE는 비뚤림 위험평가 도구를 구체적으로 제시하지 않아 적용상의 문제 있어서 평가도구를 명확히 제시함. -1 혹은 -2의 대략적인 기준을 정할 필요 있기 때문에 대략적인 판단틀을 제시함. 	

㉡ 비일관성

GRADE	한국보건의료연구원
<ul style="list-style-type: none"> • 설명되지 않는 비일관성(I², 효과추정치의 방향성과 신뢰구간의 겹침 등으로 판단)이 있으면 한 등급 혹은 두 등급 낮춘다. • 설명되면 나누어 분석한다. 	<ul style="list-style-type: none"> • 메타분석이 있는 경우 GRADE에 준해서 판단한다. • 메타분석이 없는 경우 인구 집단, 중재, 결과의 동질성으로 판단할 수 있다. • 단일 연구인 경우 판단할 수 없으며 이 경우 낮출 지 여부를 정성적으로 판단한다. • I²가 대략적으로 50-75%이면 한 등급을 낮추고 75% 이상이면 두 등급을 낮추지만 효과 크기, 방향성 등을 고려한다.
<ul style="list-style-type: none"> • 설명: 메타분석이 없는 경우 등급을 낮출 지 여부를 정성적으로 판단하며 판단 이유를 기술한다. I²에 대한 다른 기준을 적용할 경우 합당한 이유를 기술한다. 	

㉔ 비직접성

GRADE	한국보건 의료연구원
<ul style="list-style-type: none"> • 간접비교, PICO의 비직접성, 대리 결과인 경우 근거 등급 수준을 낮춘다. 	<ul style="list-style-type: none"> • GRADE에 준해서 판단하며, 추가적으로 인종, 문화적인 차이가 없다고 판단되는 경우를 제외하고는 한 등급 낮춘다. • 적용성은 판단하지 않는다. • 메타분석이 없는 경우 정성적으로 판단한다.

㉕ 비정밀

정밀도의 판단은 다음과 같은 단계를 거친다.

- 첫째. OIS(Optimal information size)를 만족하지 않으면 한 등급 또는 두 등급을 낮춘다. 그러나 각 군당 1,000명(최소한 1,000명, 가능하면 2,000명) 이상이면 낮추지 않는다.
- 둘째. OIS를 만족하면서 신뢰구간이 1을 통과하지 않으면 등급을 낮추지 않는다.
- 셋째. OIS를 만족하면서 신뢰구간이 1을 통과하는 경우 신뢰구간이 0.75 혹은 1.25를 통과하면 등급을 낮춘다.

GRADE	한국보건 의료연구원
<ul style="list-style-type: none"> • 변수의 종류에 따라서 본문의 방법대로 시행한다. 	<ul style="list-style-type: none"> • GRADE에 준해서 판단한다. • 단순화시킬 경우 반드시 통계 지문을 받아야 한다.
<ul style="list-style-type: none"> • 설명: WHO의 경우 대상자 수 100-200명이거나 RR의 신뢰구간이 1을 통과하면서 0.5 혹은 1.5를 통과하면 1 등급을 낮추고 대상자 수가 100명 이하이면서 RR의 신뢰구간이 1을 통과하면서 0.5 혹은 1.5를 통과하면 두 등급을 낮춘다. 	

㉖ 출판 비뒤림

GRADE	한국보건 의료연구원
<ul style="list-style-type: none"> • Funnel plot으로 판정하며 정도에 따라 한 등급 혹은 두 등급 낮춘다. • 출판물의 수가 적고 모든 경우 중재 이득이 있는 경우 의심 	<ul style="list-style-type: none"> • 평가하지 않는다.
<ul style="list-style-type: none"> • 설명: 메타분석에서 funnel plot을 제시하지 않는 경우가 많고, 제시하고 있더라도 평가하기 어려운 경우가 많다. AHRQ에서도 보조적인 수단으로만 판단하며, 출판 비뒤림은 일관성, 정밀도, 효과의 크기에 이미 반영된 것으로 판단한다. 	

㉔ 효과의 크기, 교란변수의 효과, 양-반응관계

효과크기	효과 측정	근거의 질
크다	RR > 2 or < 0.5	1등급 높임
매우 크다	RR > 5 or < 0.2	2등급 높임

항목	GRADE	한국보건의료연구원
효과크기	<ul style="list-style-type: none"> 위의 기준 	<ul style="list-style-type: none"> GRADE에 준해서 판단한다. 타당도에 위협이 없는(어떠한 이유든 등급을 낮추는 요인이 없는) 관찰연구만 평가
교란변수	<ul style="list-style-type: none"> 모든 교란 변수가 효과의 크기를 줄이는 쪽으로 반영할 때 관찰연구와 무작위 대조 연구 	<ul style="list-style-type: none"> GRADE에 준해서 판단한다. 무작위배정 비교임상시험과 타당도에 위협이 없는 관찰연구만 평가
양-반응관계	<ul style="list-style-type: none"> 양-반응 관계 관찰연구만(추정) 	<ul style="list-style-type: none"> GRADE에 준해서 판단한다. 타당도에 위협이 없는(어떠한 이유든 등급을 낮추는 요인이 없는) 관찰연구만 평가

㉕ 근거수준의 등급

GRADE	한국보건의료연구원
<ul style="list-style-type: none"> 높음, 중등도, 낮음, 매우 낮음으로 평가함 	<ul style="list-style-type: none"> 높음, 중등도, 낮음, 불충분(혹은 I, II, III, IV)으로 평가함
<ul style="list-style-type: none"> 설명: '매우 낮음'의 경우 실제적인 정의는 불충분에 가깝고, 국내에서 관습적으로 불충분이란 용어를 사용하고 있다. 	

㉖ 전체적 근거수준

GRADE	한국보건의료연구원
<ul style="list-style-type: none"> 근거수준이 핵심 결과에 따라 결과의 방향이 다르면: 가장 낮은 근거의 질에 따라 전체적인 근거의 질을 판단 모든 결과의 방향이 같다면: 핵심 결과의 최고 근거의 질에 따라 판단, 이득과 위해의 저울질이 불확실하면 가장 낮은 것으로 결정하여야 한다. 	<ul style="list-style-type: none"> GRADE에 준해서 판단한다.
<ul style="list-style-type: none"> 설명: GRADE와 동일한 방법으로 전체적 근거수준을 평가한다. 	

② 권고등급

㉠ 권고등급의 결정

GRADE	한국보건료연구원
<ul style="list-style-type: none"> • 근거의 질, 이득과 유해의 저울질과 가치와 선호도를 고려하여 결정 • 비용을 고려하여 문제가 있으면 다시 결정 	<ul style="list-style-type: none"> • 근거수준과 이득, 유해의 저울질로 결정 • 가치와 선호도를 참조한다.
<p>• 설명 : 국내 상황에서 가치와 선호도를 명확하게 판단하기 어렵고, 모든 권고에 비용을 고려하기는 쉽지 않다. 따라서 가치와 선호도는 명백히 문제가 있는 경우에만 권고등급에 영향을 미치며, 비용의 문제로 권고의 등급을 낮추지 않는다.</p>	

㉡ 권고등급안

GRADE	한국보건료연구원
<ul style="list-style-type: none"> • 강한 : 이득이 비용, 유해보다 명백히 크거나 비용, 유해가 명백히 이득보다 클 때 • 약한 : 근거수준이 낮거나 이득과 유해가 크기 차이가 명백하지 않음 	<ul style="list-style-type: none"> • 변경된 권고등급 결정 방법에 따라 권고등급을 결정함 • 다음과 같이 등급화 함 <ul style="list-style-type: none"> - strong for : A - weak for : B - weak against : C - strong against : D - insufficient : In

4) 다양한 경우의 GRADE 적용

원래 GRADE는 메타분석이 있는 체계적 문헌고찰의 경우에만 적용이 가능하다. 하지만 현실적으로 모든 핵심질문에 대해서 메타분석이 있는 체계적 문헌고찰이 존재하지 않고, 이런 경우 새로(*de novo*) 체계적 문헌고찰을 시행하기도 어렵기 때문에 GRADE 적용에 문제가 발생할 수 있다.

메타분석을 수행한 체계적 문헌고찰이 없는 경우 이미 개발된 진료지침 혹은 기존의 체계적 문헌고찰에서 정보를 얻을 수 있고, 이런 경우 GRADE를 다음과 같이 적용할 수 있다.

(1) 기존의 진료지침에서 정보를 얻는 경우

이미 개발된 진료지침을 근거의 출처로 하여 새로운 진료지침을 만드는 경우이며, 이런 것을 진료지침의 수용개작(adaptation)이라고 한다.

수용개작을 이용해서 진료지침을 개발하기 위해서는 다음과 같은 조건이 필요하다.

- 해당 주제에 대한 적절한 진료지침의 존재
- 중요한 최신 근거를 포함하고 있음
- 진료지침의 주제 범위가 적절함
- 진료지침의 일정 수준 이상의 질
- 해당 권고를 수용하거나 적용하기에 문제없음

만일 위의 조건을 만족하여 수용개작 과정을 거쳐 진료지침을 개발하고 여기에 GRADE 등급을 적용하기 위해서는 몇 가지 단계를 거쳐야 한다.

① 근거수준의 판단

기존 진료지침에 부여 되어 있는 근거수준의 정의와 GRADE의 근거수준 결정 방법을 비교하여 그대로 번역하여 적용할 수 있는 지 검토한다. 아래 표에서 보는 바와 같이 USPSTF와 AHRQ에서 High로 평가된 것은 GRADE에서 High로 번역해도 큰 문제가 없다(표 15). 만일 이러한 번역이 어려운 경우에는 근거 자체 혹은 근거 요약을 검토하여 정성적으로 판단할 수 있다.

표15. GRADE와 USPSTF, AHRQ의 근거수준 평가 방법

	GRADE	USPSTF	AHRQ
연구 설계	<ul style="list-style-type: none"> • RCT : High • 관찰 연구 :Low • 환자군 연구 :very low 	<ul style="list-style-type: none"> • 무작위 대조임상시험(I) • 무작위 시험이 아닌 잘 고안된 대조임상시험 (II-1) • 잘 고안된 코호트 연구나 환자 - 대조군연구 (II-2) • 중재 유무와 관계없이 여러 관찰연구(II-3) • 임상적 경험, 기술적 연구, 전문단체의 보고 (III) 	연구설계와 ROB를 평가해서 <ul style="list-style-type: none"> • Low risk of Bias • Medium • High
비뚤림 위험	특정 도구 지정없음 <ul style="list-style-type: none"> • RCT : Cochrane ROB와 유사 • NRS : 명확하지 않음 	<ul style="list-style-type: none"> • 체계적 문헌고찰인 경우 : 논문 검색 방법의 완전성, 연구 평가 기준, 결론의 타당성, 적절성 • 환자-대조군 연구 : 환자군 결정의 정확성 등 • 코호트 연구 • 무작위배정 비교임상시험 	

일관성	메타 분석의 I ² , P값 등으로 판단	<ul style="list-style-type: none"> • 응집성(cohesion)/ 일관성(consistency) • 판단에 대한 기준 없음 	주로 정성적으로 판단 Three levels of consistency: <ul style="list-style-type: none"> • Consistent (i.e. no inconsistency) • Inconsistent • Unknown or not applicable (e. g. single study)
직접성	명백한 비직접성 기준 제시	<ul style="list-style-type: none"> • 외적 타당도 • 판단에 대한 기준 없음 	GRADE와 동일
정밀도	메타분석을 통해 제시	없음	메타분석을 시행하지 않은 경우 정성적으로 판단

② 권고등급의 결정

해당 증재에 대해서 번역된 근거수준과 이득과 유해의 저울질을 바탕으로 권고등급을 결정하며, 가치와 선호도를 고려할 수 있다.

(2) 기존의 체계적 문헌고찰을 이용하는 경우

이미 발표된 체계적 문헌고찰을 이용하여 GRADE 등급을 결정하는 방법이다. 이러한 방법을 이용해서 진료지침을 개발하기 위해서는 해당 증재에 대한 적절한 체계적 문헌고찰이 존재하며, 최신성, 적절한 주제 범위, 질 등에서 문제가 없어야 한다.

이런 경우 메타분석이 있다면 한국보건의료원 GRADE 평가방법에 따라 평가를 진행하고, 메타분석이 없다면 정성적으로 해당 요소에 대해 판단해서 근거수준과 권고등급을 결정한다.

2. 권고안 도출

1) 근거수준에 대한 적절한 표현

체계적 문헌고찰을 통해 도출된 결론을 적절한 어휘로 전달하는 것은 매우 중요하다. 도출된 결론에 따른 추천되는 우리말 표현은 표 16과 같다.

표 16. 체계적 문헌고찰 결론에 따른 추천되는 우리말 표현

체계적 문헌고찰 결론	추천되는 우리말 표현	코멘트
Review found no evidence at all	<ul style="list-style-type: none"> 해당 중재에 대한 임상 연구가 없다. 	"근거가 없다"라는 표현은 오해를 일으킬 수 있음
Review found evidence that clearly supports intervention	<ul style="list-style-type: none"> 해당 중재는 사망 위험을 줄였다. 해당 중재가 효과적이라는 근거가 있다. 해당 중재는 근거가 있다. 	첫 번째 표현이 추천되며 효과의 정도(중등도 등)를 가능하면 표현하도록 한다.
Review found clear evidence of lack of benefit	<ul style="list-style-type: none"> 해당 중재는 사망 위험을 줄이지 못하였다. 해당 중재가 치료에 효과적이지 않다. 해당 중재가 치료에 효과적이라는 근거는 없다. 	첫 번째 표현이 추천된다.
Review found insufficient evidence that supports intervention	<ul style="list-style-type: none"> 근거가 부족하다. 근거가 불충분하다. 근거가 불확실하다. 	가장 표현하기 어려우며 효과가 없다는 것과 구별이 되어야함
Review found evidence of important trade-offs between known benefits and known adverse effects	<ul style="list-style-type: none"> 효과 여부에 대해 판단내리기 힘들다. 일부 환자에서 효과적이다. 근거가 불확실하다. 	근거는 불충분하지 않지만 결과의 일관성 등으로 결론을 내리기 힘든 경우이다.

2) 권고안의 도출

권고 도출의 원칙은 아래와 같다.

(1) 권고안은 특이적이며 모호하지 않아야 한다.

권고안은 근거에 기초하여 특정 상황과 환자 집단에 적절한 관리가 무엇인지에 대한 구체적이고 정확한 기술을 담고 있어야 한다. 또한 근거가 확실하지 않거나 최선의 관리가 불확실한 경우 불확실한 부분에 대해서 진료지침에 기술하여야 한다.

(2) 주요 권고안은 쉽게 확인할 수 있어야 한다.

사용자가 주요 권고안을 쉽게 발견할 수 있어야 하며, 권고안은 진료지침이 다루고 있는 핵심 질문에 대한 답변의 형태로 작성한다.

(3) 권고와 이를 지지하는 근거들이 명시적으로 연결되어 있어야 한다.

(4) 권고등급이 적절히 표현되어야 한다.

강한 권고는 "권고한다" 등의 표현으로, 약한 권고는 "고려한다" 등의 표현으로 제시하고, 권고등급과 근거수준은 권고안과 함께 제시한다.

(5) 권고의 대상이 되는 환자(또는 인구집단)와 권고 중재에 대해 가능한 자세히 명시하여야 한다.

확인 사항	예	아니오
1. 근거수준과 권고등급을 GRADE를 이용하여 체계적으로 도출하였다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 근거문 작성시 추천되는 우리말 표현을 이용하였다.	<input type="checkbox"/>	<input type="checkbox"/>
3. 권고안을 권고안 도출 원칙에 따라 작성하였다.	<input type="checkbox"/>	<input type="checkbox"/>

제 2부. 개발

10단계. 합의안 도출

요약

- 권고안 도출과정에서 이용된 합의 도출 방법을 기술한다.
- 권고안 내용에 이견이 적을 경우 비공식적 합의 방법론으로 권고를 도출할 수 있다.
- 권고안 내용에 이견이 있거나 근거가 부족할 경우 공식적 합의 방법론을 이용하여 권고안을 도출하는 것이 좋다.

진료지침은 객관적인 근거를 기반으로 하지만, 근거 해석의 차이, 이득과 유해 크기 측정의 이견, 가치와 선호도의 차이 등에 따라 집단 간, 개인 간 권고안에 의견이 일치하지 않을 수 있다. 또한 객관적 근거가 부족한 경우에 어떻게 권고안을 도출할 지에 대해 이견이 있을 수 있다. 이런 경우 진료지침 개발그룹 내에서 타당한 권고안을 도출할 수 있는 합리적인 방법이 필요하다. 이런 경우 적용할 수 있는 방법이 합의 도출 방법론이다.

합의 도출 방법에는 비공식적(**informal**) 방법과 공식적인(**formal**) 방법이 있다. 공식적인 방법에는 명목집단기법(**nominal group technique, NGT**), RAND법, 합의도출회의, 델파이법 등이 있다. 비공식적인 방법과 공식적인 방법을 구분하는 기준은 익명성, 반복성, 통제된 피드백 등의 요소가 있는 지 여부이다(표3).

특히 서로 의견이 일치하지 않는 영역이 어떤 것이며 이를 해결하기 위해 어떤 방법(예들 들어 투표, 공식적 합의 기법 등)을 사용하였는 지를 구체적이고 투명하게 기술하여야 한다.

NGT법, RAND법, 합의도출회의, 델파이법 중 어떤 방식을 채택할 것인가는 연구 상황, 전문가들의 지역적 분포, 연구자의 선호도 등 여러 가지 요인에 의해서 결정되며, 일반적으로 방법이 다르면 결과도 다르다. 하지만 어떤 방법이 명백하게 더 적절하다는 근거는 없으나 현재 진료지침 개발에 가장 많이 이용되는 공식적 합의 방법론은 NGT와 RAND법이다.

1. 합의 방법론

1) 비공식적 방법

전문가 여러 명이 모여 합의의 원칙이나 익명성 보장없이 결정에 도달하는 방법이다. 한 두 사람이 집단을 주도할 수 있다는 문제점이 있다.

2) 명목집단기법

NGT는 집단 내에서 상호작용을 구조화하는 것으로 각 참여자는 자신의 의견을 기록하고, 개인 의견을 모은 다음 하나씩 전체 위원회에서 토론하며, 각 개인의 판단은 통계적으로 모여서 집단의 판단으로 한다. 여러 사람의 아이디어와 활발한 토론을 이끌어내기 때문에 소수에 의해 토론이 지배될 가능성이 적지만 의견을 조율할 수 있는 숙련된 중재자가 필요하다.

3) 합의도출회의

미국 NIH (National Institute of health)에서 개발한 방법으로 캐나다나 영국 등에서도 진료지침 개발에 이용된다. 며칠 간 공개회의 형식으로 진행되며 위원회가 아닌 관심집단이나 전문가들이 근거를 제시하고 제시된 근거에 따라 합의를 시도한다. 위원회는 합의에 도달하려고 노력하지만 소수 의견이나 대체 의견을 존중하여 합의에 도달하지 않을 수도 있다.

4) 델파이법

의사소통이 우편을 통해서만 이루어지며, 연구자는 1차 응답에 대한 결과를 취합한 뒤 결과를 피드백한다. 그 후 새로운 결정에 대해 물어보고 그러한 과정을 여러 번 반복한다. 익명성이 보장되므로 대면토론에서 나타날 수 있는 바람직하지 못한 심리적 효과를 피할 수 있는 장점이 있다. 주로 기술평가, 교육, 우선순위 결정 등에 이용되고, 진료지침 개발에 사용되는 경우는 흔하지 않다.

5) RAND 방법

Modified NGT(혹은 Modified pelphi) 방법이라고도 하며, 진료지침 권고 도출에 많이 이용된다. 전문가 7~15명으로 그룹을 구성하고 사전에 그룹원들에게 우편으로 설문조사를 시행한다. 추후 대면회의에서 사전설문 결과를 배포하고 토론을 거친 후 최종적인 의견을 표시한다. 9점 척도에 표시된 중앙값을 기준으로 7~9점으로 합의되면 'appropriate'로 간주한다.

확인 사항	예	아니오
1. 권고안 도출에 이용한 합의 방법이 명시되어 있다.	□	□
2. 권고 내용에 이견이 있는 경우 공식적인 합의 방법론을 이용하였다.	□	□

제 3부. 최종화

11단계. 외부검토 및 갱신 계획

요약

- 진료지침 초안이 완성되면 해당분야 전문가나 이해당사자에게 초안을 보내고 그에 대한 의견을 받을 수 있다.
- 외부검토는 동료검토(peer review), 이해당사자의 의견 수렴, 사용자 사전 조사로 구분할 수 있다.
- 갱신 계획을 세운다.
- 진료지침 주제와 관련된 단체의 공식적 승인을 받는다.

1. 외부검토

진료지침 초안이 완성되면 해당분야 전문가나 영향을 받을 수 있는 이해당사자(임상 의사, 환자 등)에게 초안을 보내고 그에 대한 의견을 받을 수 있다. 이러한 과정을 외부검토라고 한다. 이해당사자에는 정책결정자, 의사결정자, 기구 대표 및 관리자 등이 포함될 수 있다. 외부검토를 동료검토(peer review), 이해당사자의 의견 수렴, 사용자 사전조사로 구분하기도 하며, 이해당사자 의견 수렴은 공청회, 웹 게시 후 의견 수렴 등 다양한 방법이 있다.

- 동료검토는 진료지침의 범위, 개발방법, 권고의 타당성에 대해 개발에 참여하지 않은 해당분야 전문가에게 평가를 받는 것이다.
- 이해당사자 의견 수렴은 진료지침 주제와 관련이 있거나 영향을 받을 수 있는 이해당사자(임상 의사, 환자, 기업, 정부 기관 등)에게 권고 내용 중 편향된 내용은 없는지, 사회적 가치가 충분히 반영되었는 지에 대해 의견을 듣는 것이다.
- 사용자 사전 조사는 진료지침을 실제로 사용할 사용자에게 진료지침 초안을 승인할 것인지, 장점과 약점은 무엇인지, 수정이 필요한 부분은 무엇인지에 대해 질문을 하는 것이다. 더불어 사용자들에게 개발 과정을 얼마나 신뢰할 수 있는지, 실제 임상에서 진료지침을 사용할 의사가 있는지, 진료지침이 실제 진료에 어떤 영향을 미칠 것인지에 대한 질문을 할 수 있다.

전문가 심사와 이해당사자 의견 수렴은 구분되어 시행되기도 하고 동시에 진행되기도 한다.

2. 갱신 계획

진료지침의 갱신은 새로운 근거를 확인하고, 새로운 근거가 진료지침을 갱신할 정도인지 판단하는 두 개의 과정이 필요하다. 갱신일정과 방법을 최종문서에 포함한다. 갱신일정은 미리 간격을 정할 수도 있고, 중요한 근거의 변화에 맞춰 갱신을 진행할 수도 있다.

새로운 근거는 체계적 문헌고찰 혹은 전문가 자문을 통해 확인할 수 있다. 새로운 근거가 진료지침을 갱신할 정도인지는 그 근거가 권고에 얼마나 강하게 영향을 줄 수 있는지에 따라 달려 있다 (예: 자원의 변화, 결과의 변화, 기술의 변화, 기존의 이득과 유해의 변화, 가치 변화).

갱신 이후 아래와 같이 결정할 수 있다.

- 전체 진료지침 폐기
- 일부 권고의 폐기
- 체계적 문헌고찰 다시 시행
- 진료지침의 타당성이 위협받지 않는 범위에서 갱신이 필요한 권고만 수정

3. 관련 단체의 공식적 승인(endorsement)

개발된 진료지침에 대해 해당 주제와 관련이 있는 전문가 단체의 공식적 승인을 받도록 한다. 관련 전문가 단체가 진료지침에 대해 공식적 승인을 하면 해당 단체의 구성원들이 진료지침을 수용할 가능성이 높아진다. 공식적 승인은 관련 조직이 진료지침의 존재를 인지하는 수준일 수도 있고, 개발된 진료지침이 조직 내부의 정책으로써 실행되기 위한 공식적 과정일 수도 있다. 한국보건 의료연구원의 공식적 승인을 받기 위해서는 성과확산팀에 의뢰하며, 연구성과검토 위원회를 거쳐 확정된다.

확인 사항	예	아니오
1. 진료지침 초안에 대한 전문가의 심사를 받았다.	<input type="checkbox"/>	<input type="checkbox"/>
2. 목표 사용자를 대상으로 개발된 진료지침에 대한 사전조사를 시행하였다.	<input type="checkbox"/>	<input type="checkbox"/>
3. 이해당사자의 의견 수렴을 받았다.	<input type="checkbox"/>	<input type="checkbox"/>
4. 진료지침 갱신 계획을 수립하였다.	<input type="checkbox"/>	<input type="checkbox"/>
5. 학회 혹은 관련 기구의 공식적 승인을 받았다.	<input type="checkbox"/>	<input type="checkbox"/>

제 3부. 최종화

12단계. 진료지침 출판

요약

- 실제 사용될 진료 상황에 맞게 진료지침을 맞춤화하여 출판한다.
- 출판할 때는 진료지침 보고지침 양식에 맞추는 것을 권장한다.

완성된 진료지침에 대해 실행 계획을 수립하고, 맞춤화 작업을 하는 것은 개발 과정의 일부분이다. 이때는 실제 진료에 미치는 영향과 조직적·문화적 맥락을 충분히 이해해야 하며, 이러한 사항들을 잘 고려해서 개발된 진료지침을 맞춤화 한다.

권고 내용이 짧고 분명한 경우 진료지침의 수용성이 높아진다. 진료지침 형식은 목표 사용자 집단에 맞는 형태가 되어야 한다. 최종 문서에 실행 계획을 넣을 지를 고려할 수 있다.

진료지침의 내용이 결정되었으면, 진료지침 작성 양식에 따라 진료지침 출판을 준비한다. ADAPTE에서 제시하는 진료지침 작성 양식은 표17과 같다. 또한 CPG지원국에서 개발한 임상진료지침 표준 보고가이드(STARIGs, STANdard Reporting Items for clinical practice Guidelines)를 참고하여 작성하는 것을 권장한다(표 18, 부록 5).

표 17. ADAPTE 진료지침 작성 양식

1. 요약
2. 서론 및 배경
3. 범위와 목적
4. 진료지침의 목표 사용자
5. 핵심질문
6. 권고
7. 권고에 대한 지지 근거 및 정보
8. 외부검토 및 자문 과정
9. 검토 및 갱신 계획
10. 알고리즘과 요약
11. 실행 시 유의사항
12. 용어 해설(친숙하지 않은 용어에 대해)
13. 진료지침을 만드는데 사용한 참고문헌
14. 원 진료지침 개발자의 승인 및 허가
15. 개발그룹 구성원 목록 및 경력, 이해관계 선언
16. 자금출처 목록
17. 부록

표 18. 임상진료지침 표준 보고가이드, STARIGs

구분/항목	번호	검토 항목	보고된 페이지
제목			
제목	1	진료지침임을 알 수 있도록 '임상진료지침', '진료지침', '권고안' 등의 용어를 사용하여 제목을 작성한다.	
개발주체			
개발단체	2	진료지침을 개발한 기관 혹은 단체 이름을 기술한다.	
개발그룹	3	개발그룹 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색), 학제(예: 내과, 소아과, 진료지침 개발과정에서 담당할 역할(예: 위원장, 실무간사)을 기술한다.	
진료지침 위원회	4	진료지침 개발 그룹 이외의 위원회/분과위원회 등을 운영한 경우 공식적인 명칭(예: 운영위원회, 자문위원회)과 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색, 학제(예: 내과, 소아과) 등을 기술한다.	
공식적 승인			
공식적 승인	5	해당 진료지침을 지지하고 받아들인다는 의사를 표명한 기관 혹은 단체가 있는 경우 이를 기술한다.	
요약			
요약	6	진료지침의 목적과 범위, 주요 권고안, 개발방법 등이 포함되어 있는 요약을 제시한다.	
목적과 범위			
배경	7	진료지침을 개발(갱신)하게 된 이유와 필요성을 기술한다.	
목적	8	진료지침의 목적을 기술한다. 목적에는 진료지침이 다루는 건강문제, 기대되는 편익(예: 우울증 증상 개선) 및 범주(예: 예방, 선별검사, 치료) 등이 포함된다.	
대상 인구집단	9	진료지침의 적용 대상이 되는 인구집단(예: 성인 우울증 환자, 제2형 당뇨병 환자)을 기술한다. 대상 인구집단의 성, 연령, 임상적 특성, 동반 질환 등이 포함될 수 있다.	
진료지침 사용자	10	진료지침을 사용할 것으로 예상되는 보건 의료서비스 제공자 그룹(예: 일차의료 의사, 정신과전문의)을 기술한다.	
의료 환경	11	진료지침이 실행될 것으로 예상되는 의료 환경(예: 일차의료, 입원진료)을 기술한다.	
방법론			
핵심질문	12	진료지침이 다루는 핵심질문을 대상 인구집단(P), 중재(I), 비교군(C), 결과(O) 등 PICO 형식에 맞추어 기술한다.	
문헌 수집 방법	13	문헌검색 방법(데이터베이스 명, 검색기간, 검색어 등)을 기술한다.	
국내 문헌 수집 방법	14	국내 문헌검색 수행 여부를 밝히고, 사용한 검색 방법(데이터베이스명, 검색기간, 검색어 등)을 기술한다.	

문헌 선택 기준	15	문헌의 포함/배제 기준을 제시한다.
비뚤림 위험 평가 방법	16	비뚤림 위험 평가에 사용한 도구명(예: Cochrane ROB 도구, AMSTAR)과 함께 평가한 방법(예: 독립적으로 2인이 평가를 기술한다).
근거종합	17	근거종합에 이용한 방법(예: 근거종합표, 메타분석)을 기술한다.
권고안 등급체계	18	진료지침에 사용된 근거 수준 및 권고안 등급체계의 등급과 의미를 기술한다.
건강 편익과 위해를 고려한 방법	19	진료지침 권고안 실행에 따른 잠재적인 위해(예: 약물 부작용)와 건강 편익(예: 삶의 질 향상)의 균형을 어떻게 고려하였는지 기술한다.
권고안 도출방법	20	권고안 도출에 사용한 방법(예: 명목집단 기법)을 기술한다.
동료검토 절차	21	진료지침 출간 이전에 시행한 동료검토 절차를 기술하고, 가능한 경우 동료검토 그룹의 인적사항(예: 이름, 소속기관명, 직위)을 제시한다.
권고와 해당근거		
권고안	22	대상 인구집단에서 특정 상황에 권장되는 행위를 제시한다.
대체 가능한 치료 방법	23	진료지침이 다루는 특정 중재의 선택 가능한 대안을 권고안 혹은 해당근거에 제시한다.
동료검토	24	동료검토를 통해 얻어진 결과/정보는 무엇이고, 진료지침 권고안 결정에 미친 영향을 기술한다.
근거의 제한점	25	권고안 도출에 사용된 근거의 제한점(예: 근거 부족, 근거의 이질성)을 기술한다.
연구개발 필요 주제 제시	26	권고안 도출시 근거가 부족한 영역, 향후 연구가 필요한 주제를 기술한다.
독립성		
개발기금	27	개발기금 출처(예: 보건복지부, 학회자체 예산)와 개발과정에 미친 영향을 기술한다.
이해관계 선언	28	개발그룹 구성원들의 진료지침과 관련된 이해관계를 다룬 방법과 처리한 결과를 기술한다.
실행 및 확산		
알고리즘	29	단계별 의사결정이 필요한 경우 알고리즘을 제시한다.
확산도구	30	진료지침 확산도구(예: 환자용 지침, 간이 진료지침, 요약본)를 부록문서로 제시한다.
갱신 계획		
갱신 계획	31	진료지침 갱신 계획 유무와 관련 계획(시점, 기준, 방법 등)을 기술한다.
관련 정보		
용어정의	32	진료지침의 올바른 적용을 위하여 해설이 필요한 용어가 있는 경우, 이에 대한 정의와 약어에 대한 정보를 기술한다.
참고문헌	33	진료지침 개발에 사용된 참고문헌을 기술한다.
부록	34	덧붙임이 필요한 정보(예: 세부적인 검색전략, 이해상충 선언 내용, 핵심질문)를 기술한다.

확인사항	예	아니오
1. 진료지침을 보고지침에 맞게 작성하였다.	□	□

11단계.
외부검토
및
갱신 계획

12단계.
진료지침
출판

부록 1.

부록1. 한국보건 의료연구원 이해상충 선언(COI) 양식

이해상충 공개서(Collision of Interest Disclosure)

연구과제명	
역할	□ 연구책임자 □ 공동연구자 □ 연구담당자 □ 기타 연구관련자()

본인은 상기 연구와 관련하여 특정기관과 연구 관련자로부터 연구결과에 영향을 미치는 지원과 제공에 대해 다음과 같이 확인하여 보고합니다.

- 지원기관으로부터 제한없이 사용할 수 있도록 연구 비용이나 교육 보조금, 연구기기, 자문 또는 사례금 형태로 1,000만 원 상당의 비용을 제공받았습니다.
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 총 금액을 기입해주세요)
- 지원기관의 지분이익이나 스톡옵션처럼 가격을 평가하기 어려운 이권을 제공받았습니다.
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 해당 이권을 기입해주세요)
- 기타 1,000만원 상당 또는 5%가 넘는 지분이익이나 이권을 제공 받았습니다.
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 총 금액을 기입해주세요)
- 지원기관에 공식/비공식적인 직함을 가지고 있습니다(예: 사장, 자문역, 고문 등).
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 직함을 기입해주세요)
- 시험대상에 대한 지적재산권을 가지고 있습니다(예: 특허, 상표권, 라이선싱, 로열티 등).
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 해당 지적재산권 내용을 기입해주세요)
- 본인의 가족(부모, 배우자, 자녀) 또는 가족이 소속된 회사에서 위에서 기술된 것과 같은 관계를 가지고 있습니다.
□ 예 □ 아니오 (만약, "예"라고 대답한 경우, 관련 관계를 기입해주세요)

본인이 확인한 모든 내용은 정확히 기술되었으며 만약 연구 진행 중에 재정적인 이해관계가 변동되는 이해상충(COI)이 생기는 경우 30일 이내에 한국보건 의료연구원 IRB에 보고하겠습니다.

(제출자)

(서명)

부록 2. 체계적 문헌고찰 질 평가도구

● A measurement tool to assess systematic reviews (AMSTAR)

1) AMSTAR checklist: 영문¹⁾

Question	Judgement	Description
<p>1. Was an 'a priori' design provided?</p> <p>The research question and inclusion criteria should be established before the conduct of the review.</p> <p><i>Note : Need to refer to a protocol, ethics approval, or pre-determined/ a priori published research objectives to score a "yes"</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	
<p>2. Was there duplicate study selection and data extraction?</p> <p>There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.</p> <p><i>Note : 2 people do study selection, 2 people do data extraction. consensus process or one person checks the other's work.</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	
<p>3. Was a comprehensive literature search performed?</p> <p>At least two electronic sources should be searched. The report must include years and databases used(e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.</p> <p><i>Note : If at least 2 sources + one supplementary strategy used, select "yes" (Cochrane register Central counts as 2 sources; a grey literature search counts as supplementary).</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	
<p>4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?</p> <p>The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.</p> <p><i>Note : If review indicates that there was a search for "grey literature" or "unpublished literature", indicate "yes". SINGLE database, dissertation, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	
<p>5. Was a list of studies (included and excluded) provided?</p> <p>A list of included and excluded studies should be provided.</p> <p><i>Note : Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no".</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	
<p>6. Were the characteristics of the included studies provided?</p> <p>In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.</p> <p><i>Note : Acceptable if not in table format as long as they are described as above.</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Can't answer <input type="checkbox"/> Not applicable	

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질
평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

7. Was the scientific quality of the included studies assessed and documented?

'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.

- Yes
- No
- Can't answer
- Not applicable

Note : Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?

The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.

- Yes
- No
- Can't answer
- Not applicable

Note : Might say something such as "the results should be interpreted with caution due to poor quality of included studies". Cannot score "yes" for this question if scored "no" for question 7.

9. Were the methods used to combine the findings of studies appropriate?

For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I²). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).

- Yes
- No
- Can't answer
- Not applicable

Note : Indicate "yes" if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

10. Was the likelihood of publication bias assessed?

An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).

- Yes
- No
- Can't answer
- Not applicable

Note : If no test values or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

11. Was the conflict of interest stated?

Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

- Yes
- No
- Can't answer
- Not applicable

Note : To get a "yes", must indicate source of funding or support for the systematic review AND for each of the included studies.

"Can't answer" is chosen when the item is relevant but not described by the authors; "not applicable" is used when the item is not relevant, such as when a meta-analysis has not been possible or was not attempted by the authors.

a The original wording for question #4: Was the status of publication (i.e., grey literature) used/not used as an exclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.

1) Shea B, Grimshaw J, Wells G, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC medical research methodology. 2007;7(1):10-17. Reproduced table. BioMed Central(Original publisher) and revised AMSTAR: Cochrane colloquium 2010

2) AMSTAR checklist : 국문

질문	판단	판단근거
1. '사전에' 체계적 문헌고찰의 계획이 수립되었는가? 고찰 수행 전에 핵심질문과 포함기준이 확립되어야 한다. 주: 프로토콜 유무, IRB 승인이나 사전에 연구목표로 출판된 연구를 참조하여 "예"라고 체크한다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	
2. 문헌선택과 자료추출을 여러 명이 수행하였는가? 적어도 두 명의 연구자에 의해 독립적으로 문헌선택과 자료추출이 수행되어야 하고, 의견 불일치를 해소한 합의 과정이 제시되어야 한다. 주: 두 사람이 문헌선택, 두 사람이 자료추출과 합의를 하거나 혹은 한 사람이 다른 사람의 자료추출을 체크하였음을 확인한다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	
3. 포괄적인 문헌검색을 하였는가? 적어도 두 개의 전자 자료원을 이용하여 검색되어야 한다. 검색연도와 데이터베이스(예: Central, EMBASE, MEDLINE), 주제어(MeSH 제시 가능)가 기술되어야 하고, 실행 가능한 검색전략이 제시되어야 한다. 최신지견, 종설, 교과서, 특성화된 연구 등록원(specialized register) 검토, 해당 분야 전문가 자문, 참고문헌 검토 등을 통해 검색이 보완하여야 한다. 주: 적어도 두 개의 전자 자료원과 하나의 보완 전략이 사용되었으면 "예"로 체크한다. [Cochrane register와 Central은 두 개의 전자자료원으로 보며, 회색문헌검색은 보완 자료원으로 본다]	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	
4. 포함기준에 출판상태(예: 회색문헌)가 사용되었는가? 출판여부에 관계없이 문헌이 검색되었는지, 출판상태와 언어 등에 따라 문헌을 배제했는지 여부가 기술되어야 한다. 주: 문헌고찰이 "회색문헌"이나 "미출판 문헌"을 검색했다는 명시가 있다면 "예"라고 체크한다. SIGLE 데이터베이스, 학위논문, 학회발표물, 임상연구등록 DB 모두를 회색문헌이라 고려할 수 있다. 전자자료원 검색과 회색과 비회색문헌이 모두 포함되었다면 회색문헌과 비출판문헌을 검색하였음을 명시해야 한다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	
5. 포함 및 배제된 연구 목록이 제시되었는가? 포함 및 배제된 연구 목록이 제시되어야 한다. 주: 배제된 연구가 참고문헌으로 있다면 인정할 수 있다. 목록의 전자링크가 깨졌다면, "아니요"로 체크한다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	
6. 포함된 연구의 특성이 제시되었는가? 개별연구의 연구대상, 종제, (중제)결과가 표 등의 형태로 제시되어야 한다. 분석된 연구의 특성(예: 연령, 인종, 성별, 사회경제적 상태, 질병상태, 이환기간, 중증도, 동반질환)이 제시되어야 한다. 주: 위에 제시된 특성들이 표 형태가 아니더라도 인정할 수 있다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니요 <input type="checkbox"/> 대답할 수 없음 <input type="checkbox"/> 적용할 수 없음	

7. 포함된 연구의 질이 평가되고 기술되었는가?

사전에 계획된 평가 방법을 제시하여야 한다. 예를 들어 효과성 평가 연구에서는 무작위 위약대조 이중 눈가림 연구만을 포함시킬 수 있고 배정순서 은폐를 포함기준으로 사용하기도 한다. 다른 연구 형태에는 특정기준이 더 적합할 수 있다.

- 예
- 아니요
- 대답할 수 없음
- 적용할 수 없음

주: 예를 들어 Jadad scale, risk of bias, 민감도 분석 또는 질평가 항목 제시같은 질평가 점수도 구나 체크리스트를 이용한 각 문헌의 질평가 결과가 제시되어 있더라도 문헌이 “낮음”, “높음”으로 평가되었다는 것을 명백하게 제시한다면 인정할 수 있다. 전체 연구들의 점수의 요약이나 범위를 제시했다면 인정할 수 없다.

8. 포함된 연구의 질은 결론을 도출하는데 적절히 사용되었는가?

방법론적 엄격성과 질평가 결과가 자료분석, 결론도출 시 고려되었다.

- 예
- 아니요
- 대답할 수 없음
- 적용할 수 없음

주: 포함된 연구들의 질이 낮아 결과는 주의깊게 해석되어야 한다” 등의 언급이 있을 수 있다. 7 번 항목에 “아니요”로 체크했다면 이 항목에서 “예”로 평가할 수 없다.

9. 개별연구의 결과를 결합한 방법이 적절하였는가?

연구들의 동질성을 평가하여 결과의 결합 가능성이 검증되어야 한데에: 동질성에 대한 카이 제곱 검정, I². 이질성이 있다면 무작위 효과 모형(random effects model)을 사용하고 결과를 결합하는 것이 임상적으로 적절한 지 고려되어야 한데에: 결합하는 것이 합리적인가?.

- 예
- 아니요
- 대답할 수 없음
- 적용할 수 없음

주: 예를 들어 중재간의 이질성/다양성으로 인해 통합하지 못했다고 설명하거나, 이질성에 대해 언급하거나 설명하였다면 “예”로 체크한다.

10. 출판 비틀림의 가능성을 평가하였는가?

출판 비틀림의 가능성을 그래프(예: funnel plot 등) 또는 통계적 검정 결과(예: Egger 회귀검정)로 평가하여야 한다.

- 예
- 아니요
- 대답할 수 없음
- 적용할 수 없음

주: unnel plot이나 검사 결과값이 포함되지 않았다면, “아니요”에 체크한다. 포함된 연구가 10 개 미만이라면 출판 비틀림이 평가되지 못했다는 언급이 있다면 “예”로 평가한다.

11. 이해상충이 기술되었는가?

체계적 문헌고찰 및 포함된 연구들의 연구비 출처가 명확하게 제시되어야 한다.

- 예
- 아니요
- 대답할 수 없음
- 적용할 수 없음

주: “예”로 평가하려면, 체계적 문헌고찰의 연구비 출처나 지원에 대해 명시해야 하며, 또한 포함된 연구 각각의 연구비 출처가 명시되어야 한다.

“대답할 수 없음”: 시행할 수 있으나 시행여부가 기술되지 않은 경우

“적용할 수 없음”: 시행할 수 없는 경우(예: 메타분석이 가능하지 않거나 저자에 의해 시도되지 않은 경우)

부록 3. 무작위배정 비교임상시험 질 평가도구

● The Cochrane Collaboration's tool for assessing risk of bias

1) Revised Cochrane Risk of Bias: 영문

RANDOM SEQUENCE GENERATION	
Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.	
Criteria for a judgement of 'Low risk' of bias	<p>The investigators describe a random component in the sequence generation process such as:</p> <ul style="list-style-type: none"> • Referring to a random number table; • Using a computer random number generator; • Coin tossing; • Shuffling cards or envelopes; • Throwing dice; • Drawing of lots; • Minimization*. <p>* Minimization may be implemented without a random element, and this is considered to be equivalent to being random.</p>
Criteria for the judgement of 'High risk' of bias	<p>The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example:</p> <ul style="list-style-type: none"> • Sequence generated by odd or even date of birth; • Sequence generated by some rule based on date (or day) of admission; • Sequence generated by some rule based on hospital or clinic record number. <p>Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example:</p> <ul style="list-style-type: none"> • Allocation by judgement of the clinician; • Allocation by preference of the participant; • Allocation based on the results of a laboratory test or a series of tests; • Allocation by availability of the intervention.
Criteria for the judgement of 'Unclear risk' of bias	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'.

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

ALLOCATION CONCEALMENT

Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.

Criteria for a judgement of 'Low risk' of bias	Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: <ul style="list-style-type: none">• Central allocation (including telephone, web-based and pharmacy-controlled randomization);• Sequentially numbered drug containers of identical appearance;• Sequentially numbered, opaque, sealed envelopes.
Criteria for the judgement of 'High risk' of bias	Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: <ul style="list-style-type: none">• Using an open random allocation schedule (e.g. a list of random numbers);• Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered);• Alternation or rotation;• Date of birth;• Case record number;• Any other explicitly unconcealed procedure.
Criteria for the judgement of 'Unclear risk' of bias	Insufficient information to permit judgement of 'Low risk' or 'High risk'. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement – for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.

BLINDING OF PARTICIPANTS AND PERSONNEL

Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.

Criteria for a judgement of 'Low risk' of bias	Any one of the following: <ul style="list-style-type: none">• No blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding;• Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias	Any one of the following: <ul style="list-style-type: none">• No blinding or incomplete blinding, and the outcome is likely to be influenced by lack of blinding;• Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken, and the outcome is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias	Any one of the following: <ul style="list-style-type: none">• Insufficient information to permit judgement of 'Low risk' or 'High risk';• The study did not address this outcome.

BLINDING OF OUTCOME ASSESSMENT

Detection bias due to knowledge of the allocated interventions by outcome assessors.

Criteria for a judgement of 'Low risk' of bias	<p>Any one of the following:</p> <ul style="list-style-type: none"> • No blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding; • Blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias	<p>Any one of the following:</p> <ul style="list-style-type: none"> • No blinding of outcome assessment, and the outcome measurement is likely to be influenced by lack of blinding; • Blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias	<p>Any one of the following:</p> <ul style="list-style-type: none"> • Insufficient information to permit judgement of 'Low risk' or 'High risk'; • The study did not address this outcome.

INCOMPLETE OUTCOME DATA

Attrition bias due to amount, nature or handling of incomplete outcome data.

Criteria for a judgement of 'Low risk' of bias	<p>Any one of the following:</p> <ul style="list-style-type: none"> • No missing outcome data; • Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); • Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; • Missing data have been imputed using appropriate methods.
Criteria for the judgement of 'High risk' of bias	<p>Any one of the following:</p> <ul style="list-style-type: none"> • Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size; • 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; • Potentially inappropriate application of simple imputation.

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

Criteria for the judgement of 'Unclear risk' of bias	Any one of the following: <ul style="list-style-type: none"> • Insufficient reporting of attrition/exclusions to permit judgement of 'Low risk' or 'High risk' (e.g. number randomized not stated, no reasons for missing data provided); • The study did not address this outcome.
--	---

SELECTIVE REPORTING

Reporting bias due to selective outcome reporting.

Criteria for a judgement of 'Low risk' of bias	Any of the following: <ul style="list-style-type: none"> • The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; • The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon).
--	--

Criteria for the judgement of 'High risk' of bias	Any one of the following: <ul style="list-style-type: none"> • Not all of the study's pre-specified primary outcomes have been reported; • One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; • One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); • One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; • The study report fails to include results for a key outcome that would be expected to have been reported for such a study.
---	--

Criteria for the judgement of 'Unclear risk' of bias	Insufficient information to permit judgement of 'Low risk' or 'High risk'. It is likely that the majority of studies will fall into this category.
--	--

OTHER BIAS

Bias due to problems not covered elsewhere in the table.

Criteria for a judgement of 'Low risk' of bias	The study appears to be free of other sources of bias.
Criteria for the judgement of 'High risk' of bias	There is at least one important risk of bias. For example, the study: <ul style="list-style-type: none"> • Had a potential source of bias related to the specific study design used; or • Has been claimed to have been fraudulent; or • Had some other problem.
Criteria for the judgement of 'Unclear risk' of bias	There may be a risk of bias, but there is either: <ul style="list-style-type: none"> • Insufficient information to assess whether an important risk of bias exists; or • Insufficient rationale or evidence that an identified problem will introduce bias.

2) Revised Cochrane Risk of Bias: 국문

무작위 배정순서 생성

무작위 순서의 부적절한 생성에 따른 선택 비돌림(중재 배정 비돌림)

비돌림 위험 '낮음' 기준	<p>순서 생성에 무작위방법을 시행한 경우 예를 들어,</p> <ul style="list-style-type: none"> • 난수표 이용 또는 컴퓨터를 이용한 난수 생성 • 동전던지기, 카드나 봉투섞기(꺼낸 카드는 다시 집어넣어야 함), 주사위 던지기, 심지뽑기, *최소화법 등 사용. 그러나 '난수'임이 보장되는 수행과정 확인할 수 있어야 적절한 방법으로 볼 수 있음. 예를 들어, 동전을 던져서 앞면이 나오면 뒷면이 나올 때까지 다시 던지지 않았음을 확인할 수 있어야 함 (배정순서 은폐와 연결됨). <p>* 최소화법은 엄밀히 말해서 무작위로 순서가 배정 되는 것이라고 볼 수는 없으나 적절한 과정에 의해 수행된 최소화법은 제대로 수행된 무작위배정으로 간주함.</p>
비돌림 위험 '높음' 기준	<p>순서 생성에 무작위방법을 시행하지 않았거나 부적절한 방법을 사용한 경우 예를 들어,</p> <ul style="list-style-type: none"> • 생년월일, 내원일 등의 규칙을 이용한 배정 • 환자 등록번호 또는 병력번호의 홀수 짝수 등 규칙을 이용한 배정 • 임상가의 판단에 따른 배정 • 환자의 선호도에 따른 배정 • 검사결과에 의한 배정 • 검사결과 순 또는 약제가 준비되는 순 등 이용가능 순에 의한 배정 • 배정자가 임의로 배정
비돌림 위험 '불확실' 기준	<p>무작위 배정순서 방법에 대한 비돌림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우</p>

배정순서 은폐

부적절한 배정순서 은폐에 따른 선택 비돌림(중재 배정 비돌림)

비돌림 위험 '낮음' 기준	<p>적절한 방법에 의해 배정순서가 은폐됨으로써 연구자가 배정내용을 알 수 없는 경우 예를 들어,</p> <ul style="list-style-type: none"> • 독립적인 중앙 무작위배정 및 관리(웹기반, 전화, 제 3의 관리기관에 의한 무작위배정 통제 등) • 무작위배정순서에 의해 일련번호가 기록되어 있는 동일한 모양의 포장 사용 • 일련번호가 기록된 불투명하고 봉해진 봉투에 의한 배정순서 보관 및 개봉
비돌림 위험 '높음' 기준	<p>배정순서가 은폐될 수 있는 방법을 사용하지 않았거나 부적절한 방법의 사용에 의해 배정순서가 은폐되지 않은 경우 예를 들어,</p> <ul style="list-style-type: none"> • 난수 또는 무작위배정순서가 기재된 표를 이용한 무작위 이행 • 밀봉되지 않거나 투명하거나 일련번호가 없는 등 안전장치가 없는 무작위배정 봉투를 사용 • 교대 혹은 순환법 등의 순서를 사용 • 생일, 병력번호 등을 이용

비뚤림 위험 '불확실' 기준	배정순서 은폐 방법에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우, 예를 들어 봉투에 의해 배정되어 있다고 했으나 일련번호, 밀봉, 투명 여부에 대한 기술이 없을 때
-----------------	---

연구 참여자, 연구자에 대한 눈가림

연구 참여자, 연구자가 배정된 중재를 알게 됨으로 인한 실행 비뚤림

비뚤림 위험 '낮음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 눈가림이 시행되지 않았거나 불완전하나, 눈가림이 [중재결과에 영향을 미치지 않을 것으로 판단되는 경우 • 눈가림을 채택하여 수행하였고 연구 참여자와 연구자에 대한 눈가림이 깨지지 않았을 것으로 확신되는 경우
----------------	--

비뚤림 위험 '높음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 연구 참여자와 연구자에 대한 눈가림이 시도되었으나 눈가림이 유지되지 않았을 것으로 판단되고, 눈가림이 결과평가에 영향을 미칠 것으로 판단되는 경우 • 눈가림이 [중재결과에 영향을 미칠 수 있는 경우임에도 눈가림을 시행하지 않았거나, 눈가림을 시도하였으나 방법이 부적절한 경우
----------------	--

비뚤림 위험 '불확실' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 눈가림에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우 • 연구에서 해당 결과를 다루지 않은 경우
-----------------	---

결과평가에 대한 눈가림

결과평가자가 배정된 중재를 알게 됨으로 인한 결과 확인 비뚤림

비뚤림 위험 '낮음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 결과평가에 대한 눈가림을 채택하여 수행하였고 결과평가자에 대한 눈가림이 깨지지 않았을 것으로 확신되는 경우 • 눈가림이 시행되지 않았으나, 눈가림이 결과평가에 영향을 미치지 않을 것으로 판단되는 경우
----------------	--

비뚤림 위험 '높음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 눈가림이 결과평가에 영향을 미칠 수 있는 경우임에도 눈가림을 시행하지 않은 경우 • 결과평가자에 대한 눈가림이 시도되었으나 눈가림이 유지되지 않았을 것으로 판단되고, 눈가림이 결과평가에 영향을 미칠 것으로 판단되는 경우
----------------	---

비뚤림 위험 '불확실' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 눈가림에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우 • 연구에서 해당 결과를 다루지 않은 경우
-----------------	---

불충분한 결과자료

불충분한 결과자료의 특성이나 처리로 인한 탈락 비뚤림

비뚤림 위험 '낮음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 결측치가 없는 경우 • 결측치가 결과에 영향을 미치지 않는 경우(생존분석에서는 결측이 절단값으로 다루어짐) • 결측치가 증재군 간에 유사하게 발생하고 결측치가 발생한 원인도 유사함 • 이분형 변수의 경우 결측치 분율이 관찰발생위험을 비추어볼 때 증재효과 추정에 임상적으로 유의한 차이를 낼 것으로 보이지 않는 경우 • 연속형 변수의 경우 결측값들로부터 예견되는 증재효과의 크기가 관찰된 효과의 크기 추정에 임상적으로 유의한 영향을 미칠 것으로 보이지 않는 경우 • 적절한 통계적 방법을 사용하여 결측치를 대체한 경우
비뚤림 위험 '높음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 상당수의 결측치가 존재하고 결측치의 원인이 실제 결과에 영향을 미칠 수 있는 경우 - 증재군 간의 불균형한 결측치 수 차이 자체 또는 결측이 생긴 이유가 결과에 비뚤림을 초래할 수 있는 경우 • 이분형 변수의 경우 결측치 분율이 결과변수의 관찰발생위험에 비추어 상당 수여서 증재효과 추정에 임상적으로 유의한 차이를 낼 것으로 보이지 않는 경우 • 연속형 변수의 경우, 결측 결과로부터 예견되는 군간 증재효과 차이가 [평균의 차이 혹은 표준화 평균의 차이]가 효과크기 추정결과에 임상적으로 유의한 비뚤림을 초래하기에 충분한 경우 • 무작위 배정된 증재를 받지 않은 사람이 상당수 임에도 증재 받은 대로만 분석을 수행하여(per-protocol analysis) 결과자료를 제시한 경우 • 부적절한 방법으로 결측치를 대체한 경우
비뚤림 위험 '불확실' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 배제/탈락에 대한 보고가 불충분한 경우(예, 무작위수 언급 없음, 결측 이유에 대한 언급 없음) • 연구에서 해당 결과를 다루지 않은 경우

선택적 보고

선택적 결과 보고로 인한 보고 비뚤림

비뚤림 위험 '낮음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 프로토콜이 존재하여 연구에서 사전에 정의해놓은 일차, 이차 (증재결과들의 정의 및 분석이 사전에 정해진 방법대로 다루어졌음을 확인할 수 있는 경우) • 프로토콜은 없지만 사전에 계획된 것을 포함하여 예상되는 모든 결과를 보고하고 있는 경우
비뚤림 위험 '높음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 사전에 정해진 (증재)결과이었음에도 결과가 보고되지 않은 것이 있는 경우 • 보고된 결과 중 사전에 정한대로 방법으로 측정하거나 분석하지 않은 경우 또는 사전에 정하지 않은 (증재)결과를 보고하는 경우 (이런 분석과 보고가 이루어진 데 대한 명백한 이유와 설명-예를 들어, 예상치 못한 부작용 등-이 있는 경우는 예외) • 불완전한 결과보고로 인해 메타분석에 포함시킬 수 없는 경우 • 현 연구에서 당연히 분석되었을 것으로 예상되는 핵심결과에 대한 보고가 없는 경우
비뚤림 위험 '불확실' 기준	<p>'높음', '낮음'에 대한 판단을 위한 정보가 충분하지 않은 경우(대다수의 연구들이 이 범주에 포함될 가능성이 있음)</p>

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질
평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

연구 참여자, 연구자에 대한 눈가림

연구 참여자, 연구자가 배정된 중재를 알게 됨으로 인한 실행 비뚤림

비뚤림 위험 '낮음' 기준	그 외 비뚤림이 없는 것으로 보임
비뚤림 위험 '높음' 기준	추가 비뚤림의 위험이 있는 것으로 판단 예를 들어, <ul style="list-style-type: none">• 특정 연구 설계와 관련된 잠재적 비뚤림 위험이 있음• 연구수행에 부정적이었다는 주장이 제기된 바 있음• 기타 다른 문제점을 가지고 있음
비뚤림 위험 '불확실' 기준	추가 비뚤림 가능성에 대한 여지가 있으나 비뚤림의 위험이 어느 정도일지 평가할만한 충분한 정보나 근거가 없는 경우

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

부록 4.

부록 4. 비무작위 연구의 질 평가도구

● Risk of Bias Assessment tool for Non-randomised Study

1) RoBANS 도구설명: 영문

1. Selection of participants

Selection bias caused by inadequate selection of participants

Criteria for judgement of 'Low risk' of bias	<p>Cohort study, Non-randomized controlled trial Intervention (exposure) and control groups are the same population group (identical institution and period), and the absence of outcomes among the study participants was confirmed at the starting timepoint of the study.</p> <p>Case-control study Case and control groups were selected from comparable population group. The case group was clearly defined and elucidated that the control group is not the patient group.</p> <p>Before-after study Participants were consecutively recruited, and the data was collected prospectively.</p>
Criteria for the judgement of 'High risk' of bias	<p>Any one of the following</p> <p>Cohort study, Non-randomized controlled trial</p> <ul style="list-style-type: none">• Intervention (exposure) and control groups were selected from different population groups. (e.g., intervention is differ by study periods or by study centers, or historical control groups were used)• The presence of outcomes among the study participants was not confirmed at the starting timepoint of the study. <p>Case-control study</p> <ul style="list-style-type: none">• Case and control groups are not the comparable population groups• The patient definition was made by self report or merge data.• The control group was not clearly confirmed for the inclusion of patients. <p>Before-after study</p> <ul style="list-style-type: none">• The control group was not recruited consecutively.
Criteria for the judgement of 'Unclear risk' of bias	<p>When it is unclear whether the selection of participants resulted in 'high risk' or 'low risk' of bias</p>

부록 1.
NECA
이해관계
선언(COI)
양식

부록 2.
체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.
무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.
비무작위
연구의
질
평가도구
RoBANS

부록 5.
임상진료
지침 표준
보고가이드
도구
STARIGs

2. Confounding variables

Selection bias caused by inadequate confirmation and consideration of confounding variable

Criteria for a judgement of 'Low risk' of bias	Any one of the following Non-randomized studies (except for before-after study) <ul style="list-style-type: none"> Major confounders were confirmed and considered adequately during the design phase (e.g., matching, participation restriction and others). Major confounders were confirmed and adjusted adequately during the analysis phase (e.g., stratification, propensity score, statistical adjustment and others). Before-after study <ul style="list-style-type: none"> Natural progression and learning effect* can be excluded when considering disease and intervention.
Criteria for the judgement of 'High risk' of bias	Any one of the following Non-randomized study (except for before-after study) <ul style="list-style-type: none"> Major confounding variables were not considered. Even though major confounding variables were confirmed, they were not considered adequately during the design and analysis phases. Before-after study <ul style="list-style-type: none"> Natural progression and learning effect are relatively clear when considering disease and intervention.
Criteria for the judgement of 'Unclear risk' of bias	When it is unclear whether the confounding variables resulted in 'high risk' or 'low risk' of bias

* Past experience improves future execution skills.

3. Measurement of intervention (exposure)

Performance bias caused by inadequate measurement of intervention (exposure)

Criteria for a judgement of 'Low risk' of bias	If intervention (exposure) methods were described by using at least more than one of the listed methods in below. <ul style="list-style-type: none"> Data were obtained from trustable sources such as medical records. Data were obtained from structured interviews.
Criteria for the judgement of 'High risk' of bias	Any one of the following. <ul style="list-style-type: none"> Data were obtained by a self report The clear case for Interviewer bias* The clear case for Recall bias**

Criteria for the judgement of 'Unclear risk' of bias	When it is unclear whether the intervention (exposure) measurement resulted in 'high risk' or 'low risk' of bias
--	--

* The investigated contents can be standardized to have effects on the results by the characteristics of investigators, this can be reduced by investigator training.

** When the respondents' degree of recall can have effects on the result.

4. Blinding for outcome assessment

Detection bias caused by inadequate blinding of outcome assessment

Criteria for judgement of 'Low risk' of bias	Any one of the following. <ul style="list-style-type: none"> • The outcome assessors was blinded • Although blinding is not present, its absence is judged to have no effect on the outcome measurement.
Criteria for judgement of 'High risk' of bias	Since blinding was not performed or incomplete, such blinding status is considered to have effects on the outcome measurement.
Criteria for judgement of 'Unclear risk' of bias	When it is unclear whether the blinding of outcome assessment resulted in 'high risk' or 'low risk' of bias

5. Incomplete outcome data

Attrition bias caused by inadequate handling of incomplete outcome data

Criteria for judgement of 'Low risk' of bias	Any one of the following. <p>Non-randomized studies (except for before-after study)</p> <ul style="list-style-type: none"> • No missing data could be observed. • The causes of having missing data are considered to have a relationship to the outcome (censoring does not create a bias in survival data) • The number of missing data was developed similarly in both of the intervention (exposure) - control groups, and the causes of the development are the similar. <p>Before-after study</p> <ul style="list-style-type: none"> • Information about the number of participants before and after the study existed and the baseline was not differ with completed and failed participants.
Criteria for judgement of 'High risk' of bias	Any one of the following. <p>Non-randomized studies (except for before-after study)</p> <ul style="list-style-type: none"> • The missing data could have effects on the outcome. It may be attributable to the differences of missing data in between the intervention (exposure)-control groups, or may be caused by the absence of measurement. <p>Before-after study</p> <ul style="list-style-type: none"> • Differences exist at the baseline of successful and failed participants.

Criteria for judgement of 'Unclear risk' of bias	When it is unclear whether the incomplete outcome data resulted in 'high risk' or 'low risk' of bias
6. Selective outcome reporting	
Reporting bias caused by selective outcome reporting	
Criteria for judgement of 'Low risk' of bias	Any one of the following. <ul style="list-style-type: none"> • The protocol was available, and pre-defined primary/secondary outcomes were described as planned. • All of the expected outcomes were included even without the protocols.
Criteria for judgement of 'High risk' of bias	Any one of the following. <ul style="list-style-type: none"> • The pre-defined primary outcomes were not fully reported. • Outcomes were not reported as defined previously. • Primary outcomes not pre-specified in the study existed (except for outcomes with clear explanation such as unexpected adverse effects). • Incomplete reporting on primary outcome of interest. • No report on important outcomes expected to be reported in the related field.
Criteria for judgement of 'Unclear risk' of bias	When it is unclear whether the selective outcome reporting resulted in 'high risk' or 'low risk' of bias.*

*Most of the studies fall under this category.

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

2) RoBANS 도구설명: 국문

1. 대상군 선정

부적절한 대상군 선정으로 인해 발생한 선택 비뚤림

비뚤림 위험 '낮음' 기준	<p>코호트 연구, 비무작위배정 비교임상시험 연구 중재(노출)군과 대조군이 동일한 인구 집단(기관과 기간이 동일)이며 연구 참여 시점에 연구 대상에서 (중재)결과가 없음을 확인하였다.</p> <p>환자-대조군 연구 환자군과 대조군을 비교할만한 인구 집단에서 선정했으며, 환자군은 명백히 정의되었고 대조군은 환자가 아니라는 사실을 명백히 하였다.</p> <p>전후 연구 대상군은 연속적(consecutive)으로 모집하였고, 자료는 전향적으로 수집하였다.</p>
비뚤림 위험 '높음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <p>코호트 연구, 비무작위배정 비교임상시험 연구</p> <ul style="list-style-type: none"> 중재(노출)군과 대조군이 서로 다른 인구집단이다. (연구기간 혹은 연구기관에 따라 중재가 달라지거나, 역사적 대조군인 경우) 연구 참여 시점에 연구 대상에서 (중재)결과가 있음을 확인하지 않았다. <p>환자-대조군 연구</p> <ul style="list-style-type: none"> 환자군과 대조군은 비교할만한 인구집단이 아니다. 환자 정의는 자가보고(self-report)이거나 연결자료(merge data)이다. 대조군은 환자가 아님을 확인하지 않았다. <p>전후 연구</p> <ul style="list-style-type: none"> 대상군을 연속적으로 모집하지 않았다. 후향적으로 자료를 수집하였다.
비뚤림 위험 '불확실' 기준	대상군 선정에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우

2. 교란변수

교란변수 확인과 고려가 부적절하여 발생한 선택 비뚤림

비뚤림 위험 '낮음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <p>전후 연구이외의 비무작위 연구</p> <ul style="list-style-type: none"> 주요 교란변수를 확인하였고 이를 디자인 단계(짜짓기, 참여제한 등)에서 적절히 고려하였다. 주요 교란변수를 확인하였고 이를 분석 단계(중화, 성향점수(propensity score), 통계적 보정 등)에서 적절하게 보정하였다. <p>전후 연구</p> <ul style="list-style-type: none"> 질병, 중재 등을 고려할 때 자연 경과, 학습효과* 등을 배제할 수 있다.
----------------	---

비뚤림 위험 '높음' 기준	다음 중 한 가지 이상에 해당되는 경우 전후 연구이외의 비무작위 연구 <ul style="list-style-type: none"> • 주요 교란 변수를 다루지 않았다. • 주요 교란변수를 확인하였지만 이를 디자인 단계나 분석 단계에서 적절히 고려하지 못하였다. 전후 연구 <ul style="list-style-type: none"> • 질병, 중재 등을 고려할 때 학습효과*, 자연 경과 등이 비교적 명확하다.
비뚤림 위험 '불확실' 기준	교란변수에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우

* 과거 경험이 추후 수행 능력을 향상시키는 것.

3. 중재(노출) 측정

부적절한 중재(노출) 측정으로 인해 발생한 실행 비뚤림

비뚤림 위험 '낮음' 기준	다음 중 한 가지 이상의 방법으로 중재(노출) 방법에 대해 기술하고 있는 경우 <ul style="list-style-type: none"> • 의무기록과 같이 믿음만한 기록에서 얻은 경우 • 구조화된 인터뷰에서 얻은 경우
비뚤림 위험 '높음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 자기응답에 의해 얻은 경우 • 조사자 비뚤림(interviewer bias)*이 명백한 경우 • 회상 비뚤림(recall bias)**이 명백한 경우
비뚤림 위험 '불확실' 기준	중재(노출)측정에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우

* 조사자의 특성에 따라 결과에 영향을 미치는 것으로 조사 내용 표준화, 조사자 훈련 등으로 줄일 수 있다.

** 응답자의 회상 정도가 결과에 영향을 미칠 수 있는 경우.

4. 결과 평가에 대한 눈가림

부적절한 결과 평가 눈가림으로 인해 발생한 결과 확인 비뚤림

비뚤림 위험 '낮음' 기준	다음 중 한 가지 이상에 해당되는 경우 <ul style="list-style-type: none"> • 결과 평가자에 대한 눈가림이 이루어진 경우 • 눈가림은 없지만 눈가림 여부가 결과 측정에 영향을 미치지 않는 것으로 판단되는 경우
비뚤림 위험 '높음' 기준	눈가림이 이루어지지 않았거나 불완전하며, 눈가림 여부가 결과 측정에 영향을 미치는 것으로 판단되는 경우
비뚤림 위험 '불확실' 기준	결과 평가 눈가림에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우

5. 불완전한 자료

불완전한 자료를 부적절하게 다루어 발생한 탈락 비뚤림

비뚤림 위험 '낮음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <p>전후 연구이외의 비무작위 연구</p> <ul style="list-style-type: none"> • 결측치가 없을 때 • 결측치가 생긴 이유가 결과와 관련성이 있을 것 같지 않은 경우(생존 자료에서 관찰 중단은 비뚤림을 일으키지 않는다) • 결측치가 중재(노출)군-대조군간 유사하게 발생하고 결측치가 생긴 이유도 유사하다. <p>전후 연구</p> <ul style="list-style-type: none"> • 중재 전후 대상자 수에 대한 정보가 있고, 탈락자와 원료자의 기저상태가 차이가 없다.
비뚤림 위험 '높음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <p>전후 연구이외의 비무작위 연구</p> <ul style="list-style-type: none"> • 결측치가 결과에 영향을 미칠 것으로 판단된다. 이는 중재(노출)군-대조군간 결측치의 차이 때문일 수도 있고, 결측의 이유 때문일 수도 있다. <p>전후 연구</p> <ul style="list-style-type: none"> • 탈락자와 원료자의 기저상태가 차이가 있다.
비뚤림 위험 '불확실' 기준	<p>불완전한 자료에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우</p>

6. 선택적 결과 보고

선택적 결과 보고 때문에 발생한 비뚤림

비뚤림 위험 '낮음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 프로토콜이 존재하고 사전에 정의한 일차, 이차(중재)결과가 정한대로 기술되어 있다. • 프로토콜은 없지만 예상되는 모든 결과를 포함하고 있다.
비뚤림 위험 '높음' 기준	<p>다음 중 한 가지 이상에 해당되는 경우</p> <ul style="list-style-type: none"> • 사전에 정의한 일차(중재)결과 중 보고되지 않은 것이 있다. • 정해지지 않은 방법으로(중재)결과 보고가 이루어졌다. • 사전에 정의되지 않은 일차(중재)결과가 있다(보고 하는 명백한 이유 설명(예를 들어 예상치 못한 부작용 같은)이 있는 경우는 예외). • 현재 리뷰에서 관심이 있는 일차(중재)결과에 대한 보고가 불완전하다. • 해당 분야 연구에서 보고될 것으로 예상되는 핵심(중재)결과에 대한 보고가 없다.
비뚤림 위험 '불확실' 기준	<p>선택적 결과 보고에 대한 비뚤림 위험이 '낮음', '높음' 중 어디에 해당하는지 불확실한 경우*</p>

*대부분의 연구들이 이 범주에 포함된다.

부록 5.

부록 5. 임상진료지침 표준 보고가이드



STARIGs

STandard Reporting Items for clinical practice Guidelines

임상진료지침 표준 보고가이드

Ver 1

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

1. 도구의 목적

보고가이드(reporting guideline)는 “과학적 방법을 사용하여 수행된 특정 연구를 보고 하는데 필요한 점검표, 도표 혹은 문서”로 정의된다[Moher 등(2010)]. 임상진료지침은 체계적 방법론을 적용하여 얻어진 연구 성과물로서 임상연구, 체계적 문헌고찰 등과 같이 개발방법과 주요 결과에 대하여 정확하고 투명하게 보고하는 것이 타당하다.

따라서 CPG 지원국은 진료지침 보고 수준을 개선하기 위하여 임상진료지침 표준 보고 가이드(STARIGs, STandard Reporting Items for clinical practice Guidelines)를 개발하였다. 이 도구는 진료지침 개발자가 진료지침을 충실하게 작성할 수 있도록 도와주는 체크리스트이며, 진료지침의 질 개선에 긍정적 영향을 미칠 것이다. 또한 진료지침 개발자를 위한 교육 자료로도 활용할 수 있을 것이다.

2. 임상진료지침 표준 보고가이드 구성과 내용

STARIGs는 34개 항목 및 검토항목으로 구성되었으며, 관련 내용은 다음과 같다.

- 제목 항목은 진료지침인지 파악할 수 있도록 제목을 작성하는 것에 초점을 두었음.
- 개발주체에는 3개 항목이 있으며, 개발단체의 명칭을 정확하게 기록하고, 개발그룹 특성과 함께 위원회가 구성된 경우에는 이를 기술하도록 요구하는 내용임.
- 공식적 승인 항목은 해당 진료지침을 승인한 단체나 기관을 제시해야 한다는 것임.
- 요약 항목은 독자가 해당 진료지침의 주제, 개발방법, 주요 결과를 용이하게 파악할 수 있도록 제시하는 것이 필요하다는 것임.
- 목적과 범위는 5개 항목으로 개발배경, 목적, 대상 인구집단, 진료지침 사용자, 의료 환경 등에 대해 정확히 기술해야 한다는 내용을 다루고 있음.
- 방법론은 10개 항목으로 핵심질문, 문헌 수집 방법, 국내 문헌 수집 방법, 문헌 선택 기준, 비뮌 위험 평가 방법, 근거종합, 권고안 등급체계, 건강 편익과 위해를 고려한 방법, 권고안 도출방법, 동료검토 절차에 대한 기술을 다루고 있음.
- 권고와 해당 근거에는 5개 항목이 있으며, 권고안은 특정 상황에 권고되는 행위가 기술되어야 한다는 것을 강조하고 있으며, 대체 가능한 치료 방법에 대한 기술, 동료

검토가 권고안에 미친 영향, 근거의 제한점, 향후 연구개발이 필요한 주제를 기술하도록
요구하고 있음.

- 독립성은 2개 항목으로 구성되어 있으며, 개발기금 출처와 개발과정에서 미친 영향
기술하도록 하고, 개발그룹 구성원의 진료지침과 관련된 이해관계를 밝히도록 요구하고
있음.
- 실행 및 확산은 2개 항목으로 구성되어 있으며 필요한 경우 단계별 의사결정 흐름도를
제시하고, 진료지침 확산 및 실행에 필요한 도구(환자용 지침, 간이 지침, 요약본 등)를
제시할 것을 요구하는 내용임.
- 갱신 계획은 계획이 있는지를 밝히고, 갱신 시점과 방법에 대한 내용을 기술하도록
요구하고 있음.
- 관련 정보에는 용어 정의, 참고문헌, 부록 등 3개 항목이 포함되어 있음.

3. 이용자가이드

검토항목에 대한 구체적인 내용은 이용자 가이드에 기술되어 있다. 검토항목에 대한 이해를
높이기 위하여 설명과 사례로 구성하였으며, 설명에는 검토항목의 의미, 기술할 위치에 대한
정보를 정리하였다. 사례에는 우리나라 진료지침과 외국 진료지침에서 적합하다고 판단되는
내용을 부분적으로 발췌하여 제시하였다.

4. 고려 사항

이 도구는 진료지침 개발자가 진료지침을 작성할 때 기술되어야 하는 내용을 빠짐없이
정리하였는지 확인할 수 있는 점검표(checklist)로써 진료지침 평가를 목적으로 하지는
않는다.

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질 평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

5. 참고문헌

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analysis: The PRISMA Statement. *J Clin Epidemiol*. 2009 Oct;62(10):1006-12. Epub 2009 Jul 23.

Mohr D, Schulz KF, Simera I, Altamn DG. 2010. Guidance for Developers of Health Research Reporting Guidelines. *PLoS Med* 7(2): e1000217. doi:10.1371/journal.pmed.1000217.5

NICE. The NICE Guideline on the treatment and mangement of depression on adults. NICE, 2009.

NICE. Unstable Angina and NSTEMI: the early management of unstable angina and non-ST-segment- elevation myocardial infarction. NICE, 2010.

North American Menopause Society Estrogen and progestogen use in postmenopausal women: 2010 position statement of The North American Menopause Society. *Menopause* 2010;17(2):242-255.

Plint AC, Moher D, Morrison A. Schulz K, Altman DG, Hill C, Gaboury I. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006 Sep 4;185(5):263-7.

Shaneyfelt TM, Mayo-Smith MF Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA* 1999;281(20):1900-5.

Shelley, W, McCready, D, Holloway, C, Trudeau, M, Sinclair, S and the Breast Cancer Disease Site Group. Guideline Management of Ductal Carcinoma In Situ of the Breast Cancer. *Cancer care ontario Practice Guideline Initiative*, 2006.

Shiffman RN, Shekelle P. Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Annals of Internal Medicine*. 2003;139(6):493-8.

SIGN. Management of diabetes: a national guideline. SIGN, 2010.

Item 1 제목

진료지침인지 알 수 있도록 ‘임상진료지침’, ‘진료지침’, ‘권고안’ 등의 용어를 사용하여 제목을 작성한다.

개발자는 출간물이 임상진료지침임을 쉽게 알 수 있도록 제목에 ‘임상진료지침’, ‘진료지침’, ‘권고안’ 의 용어를 사용한다. 이와 같은 용어를 사용하면 독자들이 출간물의 성격을 잘 인지할 수 있으며, 문헌 검색시 쉽게 찾을 수 있다. 또한 독자들이 쉽게 진료지침의 내용을 알 수 있도록 주요 정보를 제목에 넣는 것이 바람직하다. 즉, 진료지침이 다루는 임상적 문제(예: 우울증, 뇌졸중), 범주(예: 진단, 치료) 등이 제목에 포함되면 독자들에게 보다 분명한 정보를 줄 수 있다.

Item 2 개발단체

진료지침을 개발한 기관 혹은 단체 이름을 기술한다.

진료지침 개발을 주도적으로 진행하였거나 총체적인 책임을 맡고 있는 기관 또는 단체의 이름을 제시한다. 이 항목을 통해 독자들은 개발단체의 특성을 파악할 수 있고, 진료지침의 신뢰성, 범용성, 적용성 등을 가늠할 수 있다. 독자와 소통하기 위해 개발단체의 주소, 전화번호, 웹사이트 등을 제공하는 것이 바람직하다. 이러한 정보는 표지에 제공한다.

Item 3 개발그룹

개발그룹 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색), 학제(예: 내과, 소아과), 진료지침 개발과정에서 담당한 역할(예: 위원장, 실무간사)을 기술한다.

진료지침 개발과정에 참여한 전문가들의 정보를 제공한다. 개발그룹은 관련 임상 전문가, 방법론 전문가가 참여하며, 정보검색 전문가, 행정인력들이 참여하기도 한다. 개발그룹 구성원 중 위원장, 실무간사와 같은 특정 역할을 담당할 경우, 그 내용을 기술한다. 또한 개발그룹의 인적사항을 제시한다. 이 항목을 통해 독자들은 개발그룹의 다학제성, 학제의 적합성 등을 파악하여 진료지침의 타당성을 판단할 수 있다. 이와 같은 정보는 서문, 진료지침 개발방법을 기술한 단락 혹은 부록에 제시할 수 있다.

Item 4 진료지침 위원회

진료지침 개발그룹 이외의 위원회/분과위원회 등을 운영한 경우 공식적인 명칭(예: 운영위원회, 자문위원회)과 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색), 학제(예: 내과, 소아과) 등을 기술한다.

진료지침을 개발할 때 개발그룹이 아닌 별도의 위원회를 운영할 수 있다. 이런 위원회에는 진료지침 전반을 기획, 운영, 관리하는 운영위원회와 진료지침 내용을 검토하고 자문하는 자문위원회 등이 있다. 이 항목을 통해 독자들은 진료지침 개발 과정이 체계적 혹은 공식적으로 진행되었음을 알 수 있다. 이와 관련된 정보를 개발그룹과 같은 방식으로 제시한다.

Item 5 공식적 승인

해당 진료지침을 지지하고 받아들인다는 의사를 표명한 기관 혹은 단체가 있는 경우 이를 기술한다.

개발된 진료지침 주제와 밀접한 전문가 단체 혹은 진료지침과 관련있는 기관의 공식적 승인을 받도록 권장한다 (예: 당뇨병 진료지침- 대한가정의학회, 대한내분비학회, 대한당뇨병학회 등). 관련 전문가 단체 혹은 기관의 공식적 승인을 받으면 해당 단체의 구성원들의 진료지침 사용 가능성이 높아지고, 진료지침의 공식성, 이용 가능성이 높아진다. 공식적 승인을 받은 기관 혹은 단체명은 표지, 서문, 개발방법 단락에 기술한다.

Item 6 요약

진료지침의 목적과 범위, 주요 권고안, 개발방법 등이 포함되어 있는 요약을 제시한다.

개발자는 진료지침의 목적, 범위, 개발방법, 주요 권고안 등 핵심 정보를 요약으로 제공할 수 있다. 이를 통해 독자들은 관심있는 진료지침의 핵심 정보를 쉽게 얻을 수 있다. 통상적으로 요약은 서문과 본문 사이의 독립된 장으로 제시하고 혹은 별도 문서로 제시한다

Item 7 배경

진료지침을 개발(갱신)하게 된 이유와 필요성을 기술한다.

독자들은 진료지침 개발배경 및 이유를 이해할 필요가 있다. 여러 관점(예: 문제의 크기, 문제의 심각성, 사회적 관심)에서 해당 진료지침이 개발되어야 하는 이유를 밝힌다. 진료지침을 갱신한 경우 갱신이 필요한 이유와 함께 이전 진료지침과 비교하여 새로 추가, 변경된 내용에 대한 정보를 명시하는 것이 좋다. 이러한 정보는 서문에 제시한다.

Item 8 목적

진료지침의 목적을 기술한다. 목적에는 진료지침이 다루는 건강문제, 기대되는 편익(예: 우울증 증상 개선) 및 범주(예: 예방, 선별검사, 치료) 등이 포함된다.

진료지침이 다루는 건강문제, 기대되는 편익(예: 우울증 증상 개선) 및 범주(예: 예방, 선별검사, 치료) 등을 포함하여 목적을 기술한다. 또한 사회, 그리고 환자나 개개인의 인구집단에 미치는 진료지침의 잠재적 건강영향도 함께 기술하는 것이 바람직하다. 이 항목을 통해 독자들은 해당 진료지침이 관심있는 임상주제와 관련성이 있는지 파악할 수 있다. 목적은 배경 또는 서문에 기술한다.

Item 9 대상 인구집단

진료지침의 적용 대상이 되는 인구집단(예: 성인 우울증 환자, 제2형 당뇨병 환자)을 기술한다. 대상 인구집단의 성, 연령, 임상적 특성, 동반질환 등이 포함될 수 있다.

진료지침에서 다루는 인구집단에 대해 기술한다. 대상 인구집단의 연령, 성, 임상 상태, 질병의 중증도, 동반질환 등에 대한 정보를 제공하며, 포함하지 않는 인구집단의 특성도 같이 제시한다. 이를 통해 독자들은 관심 환자그룹과 해당 진료지침이 다루는 인구집단이 부합하는 지 판단할 수 있다. 이러한 정보는 서문에 기술한다.

Item 10 진료지침 사용자

진료지침을 사용할 것으로 예상되는 보건의료서비스 제공자 그룹(예: 일차진료의사, 정신과전문의)을 기술한다.

개발자는 진료지침의 목표 사용자를 분명하게 기술한다. 또한 목표 사용자가 해당 진료지침을 어떻게 사용(예: 임상 결정, 진료 표준에 대한 정보)할 수 있는지에 대한 정보를 제공하는 것이 바람직하다. 이를 통해 독자들은 해당 진료지침의 활용 가능성을 판단할 수 있다. 이러한 정보는 서문에 기술한다.

Item 11 의료환경(Care Settings)

진료지침이 실행될 것으로 예상되는 의료 환경(예: 일차의료, 입원진료)을 기술한다.

진료지침의 권고안에 포함된 중재가 실행될 것으로 예상되는 의료 환경 즉, 일차의료, 이차의료, 삼차의료 혹은 입원진료, 외래진료 등을 기술한다. 이를 통해 독자들은 해당 진료지침의 활용 가능성을 판단할 수 있다. 이러한 정보는 서문에 기술한다.

Item 12 핵심질문(Key questions)

진료지침이 다루는 핵심질문을 대상 인구집단(P), 중재(I), 비교군(C), 결과(O) 등 PICO 형식에 맞추어 기술한다.

개발자는 진료지침의 특정 임상영역을 핵심질문으로 구체화하고, 핵심질문은 명확하게 대답할 수 있도록 구체적이고 구조화된 형태로 구성한다. 독자들이 핵심질문을 통해 진료지침의 범위와 적용 가능성을 쉽게 파악할 수 있다. 핵심질문은 PICO를 포함하여 작성한다. 핵심질문은 서문, 개발방법 단락 혹은 부록에 제시한다.

- P (Patient/Problem): 진료지침 대상 인구집단(질병 종류, 질병 중증도, 성별, 연령 등) 특성을 구체적으로 기술한다.
- I (Intervention/Exposure): 치료방법, 진단법, 노출 등을 정의한다. 치료(노출)의 종류를 구체적으로 나타내야 하고 약물치료의 경우 용량, 용법, 투여시기, 투여경로 등을 명시한다.
- C (Comparison): 비교하는 대조군 여부를 밝히고, 대조군이 있는 경우 위약군인지 다른 중재군인지 기술한다.
- O (Outcome): 건강 결과를 기술한다(중기 또는 단기 효과측정, 사망률, 유병률, 합병증, 재발률, 재입원율, 직장복귀, 신체적 사회적 기능, 삶의 질, 전반적인 건강수준, 비용 등).

Item 13 문헌 수집 방법

문헌 검색 방법(데이터베이스명, 검색기간, 검색어 등)을 기술한다.

문헌검색에 활용한 데이터베이스별로 검색전략(예: 검색어, 검색기간, 언어 등의 제한)을 재현할 수 있을 만큼 상세하게 기술한다. 자료원은 전자 데이터베이스(예: MEDLINE, EMBASE, CINAHL), 체계적 문헌고찰 데이터베이스(예: the Cochrane Library, DARE), 진료지침 제공 웹사이트(예: the US National Guideline Clearinghouse, the German Guidelines Clearinghouse) 등이 포함된다. 이를 통해 독자들은 진료지침 범위/핵심질문에 부합되게 문헌검색이 수행되었는지 판단할 수 있다. 검색전략은 진료지침 개발방법을 기술한 장에 기술하며, 필요한 경우 별도의 문서나 부록에 기술한다.

Item 14 국내 문헌 수집 방법

국내 문헌검색 수행 여부를 밝히고, 사용한 검색 방법(데이터베이스명, 검색기간, 검색어 등)을 기술한다.

국내 문헌검색 수행 여부를 밝히고, 문헌검색에 활용한 국내 데이터베이스명, 검색기간, 검색어를 포함한 검색전략을 기술한다. 또한 전자 데이터베이스에 수록되지 않은 국내 문헌을 수기 검색한 경우에는 수행한 방법을 기술한다. 이를 통해 독자들은 권고안 도출하는 과정에서 국내 연구결과가 반영되었는지 파악할 수 있다. 또한 권고안의 수용성 혹은 적용 가능성을 판단하는데 도움이 된다. 이러한 정보는 개발방법 단락에 기술한다.

Item 15 문헌 선택 기준

문헌의 포함/배제 기준을 제시한다.

개발자는 검색에 의해 확보된 문헌의 선택 기준을 제시한다. 선택 기준에는 대상 인구집단 특성, 연구설계, 비교, 결과, 언어, 맥락 등을 포함하고, 필요시 배제 기준을 제시한다. 이 항목은 핵심질문에 적절한 문헌이 선택되었는지에 대한 정보를 제공한다. 이러한 정보는 개발방법 단락에 기술하고, 필요한 경우 별도의 문서나 부록에 기술한다.

Item 16 비뿔림 위험 평가 방법

비뿔림 위험 평가에 사용한 도구명(예: Cochrane ROB 도구, AMSTAR)과 함께 평가한 방법(예: 독립적으로 2인이 평가)을 기술한다.

비뿔림 위험이란 개별 연구에서 중재효과치 추정의 내적타당도를 의미한다. 비뿔림 위험을 평가하는 이유는 비뿔림 위험 정도에 따라 효과의 크기가 달라질 수 있기 때문이다. 문헌을 평가할 수 있는 도구는 종류가 많고, 문헌의 특성에 따라 이용되는 도구도 달라진다. 사용한 도구명과 평가 수행 방법(예: 2인이 독립적으로 평가)을 기술한다. 비뿔림 위험 평가 방법은 개발방법 단락 혹은 별도의 문서에 기술한다.

Item 17 근거종합

근거종합에 이용한 방법(예: 근거종합표, 메타분석)을 기술한다.

개발자는 근거를 종합한 방법에 대하여 기술한다. 특정 임상질문에 대해서 다양한 유형 및 수준의 연구가 존재하고, 연구 결과도 서로 일치하지 않을 수 있다. 근거를 어떻게 요약하고 종합하였는지 기술하고, 사용한 도구가 있다면 이를 기술하는 것이 바람직하다. 성인 우울증 치료 및 관리 진료지침(2009)에서는 근거 요약을 위해서는 GRADE profiler 이용하였고, 가능한 경우 정량적으로 근거를 종합할 수 있는 방법인 메타분석을 수행하였다고 제시하고 있다. 이러한 정보는 개발방법 단락 혹은 별도의 문서에 기술한다.

Item 18 권고안 등급 체계

진료지침에 사용된 근거수준 및 권고안 등급체계의 등급과 의미를 기술한다.

개발자는 진료지침에 사용한 근거수준 및 권고안 등급체계의 등급과 의미를 기술한다. 독자는 근거수준 및 권고안 등급체계의 등급과 의미를 통해서 관심있는 권고안의 근거의 양, 강도, 효과의 크기 등을 쉽게 판단할 수 있다. 이러한 정보는 진료지침 목차 앞에 제시하는 경우가 많다.

Item 19 건강 편익과 위해를 고려한 방법

진료지침 권고안 실행에 따른 잠재적인 위해(예: 약물 부작용)와 건강 편익(예: 삶의 질 향상)의 균형을 어떻게 고려하였는지 기술한다.

권고안을 도출할 때는 사망률 저하, 삶의 질 향상과 같은 건강의 편익과 약물 부작용 등의 위해를 제시하고, 이들을 어떻게 고려하였는지 기술한다. 이를 통해 도출된 권고안이 편익은 물론 위해를 충분히 고려한 균형된 형태인지를 판단할 수 있다. 이러한 정보는 개발방법 단락 혹은 권고에 기술한다.

Item 20 권고안 도출 방법

권고안 도출에 사용한 방법(예: 명목집단 기법)을 기술한다.

권고안을 도출한 방법과 최종 결정에 어떻게 도달했는지를 기술한다. 권고안 도출과정, 얻어진 결과, 도출과정이 권고안에 미친 영향에 대하여 기술한다. 예를 들면, 방법에는 투표 체계(예: 전원동의, 투표), 비공식적 합의(예: 자문 회의, 연구진 회의), 공식적 합의 기법(예: Delphi, Glaser techniques)이 포함된다. 합의를 이루지 못한 경우 이를 처리한 방법을 명시한다. 이를 통해 권고안 도출과정의 합리성, 권고안의 합의 수준을 파악할 수 있다. 이러한 정보는 개발방법 단락 혹은 별도 문서, 부록에 기술한다.

Item 21 동료검토 절차

진료지침 출간 이전에 시행한 동료검토 절차를 기술하고, 가능한 경우 동료검토 그룹의 인적사항(예: 이름, 소속기관명, 직위)을 제시한다.

일반적으로 동료검토는 진료지침 내용의 타당성을 검증하는 절차로, 진료지침을 출판하기 전에 동료검토를 수행한다. 검토자는 개발그룹에 포함되지 않아야 하고, 방법론 전문가와 임상 전문가가 수행한다. 동료검토의 목적(예: 질 개선, 권고 초안에 대한 피드백, 실행 가능성 평가)과 수행방법을 기술하고, 가능한 경우 동료검토자들에 대한 정보(예: 검토자 수, 전문성, 소속기관)를 기술한다. 이러한 정보는 개발방법 단락 혹은 별도의 문서, 부록에 기술한다.

Item 22 권고안

대상 인구집단에서 특정 상황에 권장되는 행위를 제시한다.

권고안은 대상 인구집단의 특정 상황에 권장되는 행위에 대한 기술이며, 근거 요약 또는 질병의 역학적 특성, 위험 인자 등은 해당하지 않는다. 이에 대하여 구체적이고 정확하게 기술해야 한다. 또한 권장한 행위의 목적(예: 삶의 질 개선, 부작용의 감소)을 명시한다. 이러한 정보는 권고안과 요약에 기술하고, 별도 문서를 추가로 제작할 수 있다.

Item 23 대체 가능한 치료 방법

진료지침에는 특정 상황에서 주로 권장되는 행위 이외에 선택할 수 있는 다른 중재를 권고안 혹은 해당 근거에 함께 제시한다.

진료지침에는 특정 상황에서 주로 권장되는 행위 이외에 선택할 수 있는 다른 중재(예방, 진단, 치료방법)를 함께 제시하는 것이 필요하다. 더불어 각 중재별 해당되는 인구집단의 특성 혹은 임상상황에 대해서 기술한다.

Item 24 동료검토

동료검토를 통해 얻어진 결과/정보는 무엇이고, 진료지침 권고안 결정에 미친 영향을 기술한다.

진료지침 출간 전에 수행한 동료검토 결과와 정보를 기술하고, 그 내용이 권고안에 미친 영향을 기술한다. 변경된 권고 내용과 그 이유를 공개함으로써 권고안 도출 과정의 투명성이 높아지게 된다. 이러한 정보는 권고 혹은 부록에 기술한다.

Item 25 근거의 제한점

권고안 도출에 사용된 근거의 제한점(예: 근거 부족, 근거의 이질성)을 기술한다.

검토한 근거의 양과 질에 있어서 제한점과 이에 대한 해석을 기술한다. '근거의 양이 부족했다', '이질성이 높아 메타 분석을 수행할 수 없었다', '선택된 문헌들은 심각한 비뚤림이 있다' 등과 같은 근거의 제한점이 기술되어 있으면, 독자들은 권고안을 받아들이고 실행할 때 보다 명백한 의사결정을 할 수 있다. 이러한 정보는 개발방법 단락 혹은 권고에 기술한다.

Item 26 연구개발 필요 주제 제시

권고안 도출시 근거가 부족한 영역, 향후 연구가 필요한 주제를 기술한다.

진료지침 개발과정에서 권고안 도출에 관련 근거가 부족한 부분이 파악될 수 있다. 특정 임상 영역에서 연구 및 개발이 필요한 주제를 제시해주는 것은 진료지침의 중요한 기능이다. 이를 통해 연구자들에게 필요한 연구주제에 대한 정보를 줄 수 있고, 새로운 연구비 할당에도 영향을 미칠 수 있다. 이러한 정보는 해당 근거 혹은 별도의 장에 기술한다.

Item 27 개발기금

개발기금 출처(예: 보건복지부, 학회자체 예산)와 개발과정에 미친 영향을 기술한다.

진료지침은 많은 경우 개발기금을 지원받아 개발된다(예: 정부, 전문가 단체). 개발기금 전체 지원 혹은 일부(예: 진료지침 인쇄) 재정적 지원만으로 국한될 수 있다. 진료지침 개발자는 개발기금의 출처와 그 역할을 분명하게 밝히고, 외부 기금 주체의 이익과 관점이 진료지침 내용에 미친 영향에 대하여 명시해야 한다. 이러한 정보는 표지 혹은 서문에 기술한다.

Item 28 이해관계 선언[Conflict of interest]

개발그룹 구성원들의 진료지침과 관련된 이해관계를 다른 방법과 처리한 결과를 기술한다.

진료지침 개발그룹의 모든 구성원들은 이해관계를 선언해야 한다. 이해관계 선언이란 고려한 이해관계의 유형(범주)을 밝히고, 잠재적인 이해관계를 최소화하기 위한 노력을 밝히는 것이다. 이해관계의 유형은 소속기관, 금전적 이해관계(예: 자문비, 연구비, 주식 소유), 비금전적 이해관계(예: 소속 학회, 개인적 관심), 가족 관련 이해관계 등이 있다. 또한 이해관계가 미치는 영향을 최소화하기 위해 취한 조치를 기술한다. 이러한 정보는 개발방법 단락 혹은 부록에 기술한다.

Item 29 알고리즘

단계별 의사결정이 필요한 경우 알고리즘을 제시한다.

알고리즘이란 권고 내용을 단계적 의사결정 형태의 그림으로 제시하는 것이다. 모든 진료지침에서 제시할 필요는 없으나, 의사결정이 단계적으로 이루어지는 경우 알고리즘으로 제시하면 독자들에게 훨씬 효과적으로 권고안을 전달할 수 있다. 이러한 정보는 권고 혹은 부록, 별도 문서로 제시한다.

Item 30 확산도구

진료지침 확산도구(예: 환자용 지침, 간이 진료지침, 요약본)를 부록문서로 제시한다.

권고안이 효과적으로 확산 및 실행되기 위해서는 진료지침 요약본, 간이 진료지침, reminders 등의 도구가 필요하다. 확산도구는 진료지침 부록, 별도 문서 혹은 웹사이트에서 제공될 수 있다.

Item 31 갱신 계획

진료지침 갱신 계획 유무와 관련 계획(시점, 기준, 방법 등)을 기술한다.

진료지침은 최신의 연구를 반영해야 하기 때문에 개발 이후 새로운 근거를 확인하는 과정이 필요하다. 진료지침 개발자는 해당 진료지침의 갱신 계획이 있는지 진술하고 분명한 시간 간격이나 갱신 시점을 결정할 수 있는 명료한 기준, 갱신 시점에 대한 정보를 제공해야 한다. 진료지침 갱신 계획은 진료지침 개발과정에서 결정되어야 한다. 관련한 계획은 서론, 개발방법 단락, 마감 혹은 부록에 제시한다.

Item 32 용어 정의(Glossary and abbreviations)

진료지침의 올바른 적용을 위하여 해설이 필요한 용어가 있는 경우 이에 대한 정의와 약어에 대한 정보를 기술한다.

Item 33 참고문헌(Reference)

진료지침 개발에 사용된 참고문헌을 기술한다.

진료지침 초안을 만들는데 사용된 모든 문서는 최종 문서에 참고문헌으로 정리해야 한다.

Item 34 부록(Appendix)

덧붙임이 필요한 정보(예: 세부적인 검색전략, 이해관계 선언 내용, 핵심질문)를 기술한다.

STARIGs [STAndard Reporting Items for clinical practice Guidelines]

임상진료지침 표준 보고가이드

구분/항목	번호	검토 항목	보고된 페이지
제목			
제목	1	진료지침임을 알 수 있도록 '임상진료지침', '진료지침', '권고안' 등의 용어를 사용하여 제목을 작성한다.	
개발주체			
개발단체	2	진료지침을 개발한 기관 혹은 단체 이름을 기술한다.	
개발그룹	3	개발그룹 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색), 학제(예: 내과, 소아과), 진료지침 개발과정에서 담당한 역할(예: 위원장, 실무간사)을 기술한다.	
진료지침 위원회	4	진료지침 개발 그룹 이외의 위원회/분과위원회 등을 운영한 경우 공식적인 명칭(예: 운영위원회, 자문위원회)과 구성원의 이름, 소속, 전문성(예: 방법론 전문가, 정보검색), 학제(예: 내과, 소아과) 등을 기술한다.	
공식적 승인			
공식적 승인	5	해당 진료지침을 지지하고 받아들인다는 의사를 표명한 기관 혹은 단체가 있는 경우 이를 기술한다.	
요약			
요약	6	진료지침의 목적과 범위, 주요 권고안, 개발방법 등이 포함되어 있는 요약을 제시한다.	
목적과 범위			
배경	7	진료지침을 개발(갱신)하게 된 이유와 필요성을 기술한다.	
목적	8	진료지침의 목적을 기술한다. 목적에는 진료지침이 다루는 건강문제, 기대되는 편익(예: 우울증 증상 개선) 및 범주(예: 예방, 선별검사, 치료) 등이 포함된다.	
대상 인구집단	9	진료지침의 적용 대상이 되는 인구집단(예: 성인 우울증 환자, 제2형 당뇨병 환자)을 기술한다. 대상 인구집단의 성, 연령, 임상적 특성, 동반 질환 등이 포함될 수 있다.	
진료지침 사용자	10	진료지침을 사용할 것으로 예상되는 보건의료서비스 제공자 그룹(예: 일차진료 의사, 정신과전문의)을 기술한다.	
의료 환경	11	진료지침이 실행될 것으로 예상되는 의료 환경(예: 일차의료, 입원진료)을 기술한다.	
방법론			
핵심질문	12	진료지침이 다루는 핵심질문을 대상 인구집단(P), 중재(I), 비교군(C), 결과(O) 등 PICO 형식에 맞추어 기술한다.	
문헌 수집 방법	13	문헌검색 방법(데이터베이스 명, 검색기간, 검색어 등을) 기술한다.	
국내 문헌 수집 방법	14	국내 문헌검색 수행 여부를 밝히고, 사용한 검색 방법(데이터베이스명, 검색기간, 검색어 등을)을 기술한다.	
문헌 선택 기준	15	문헌의 포함/배제 기준을 제시한다.	
비뚤림 위험 평가 방법	16	비뚤림 위험 평가에 사용한 도구명(예: Cochrane ROB 도구, AMSTAR)과 함께 평가한 방법(예: 독립적으로 2인이 평가를)을 기술한다.	
근거종합	17	근거종합에 이용한 방법(예: 근거종합표, 메타분석)을 기술한다.	
권고안 등급체계	18	진료지침에 사용된 근거 수준 및 권고안 등급체계의 등급과 의미를 기술한다.	
건강 편익과 위해를 고려한 방법	19	진료지침 권고안 실행에 따른 잠재적인 위해(예: 약물 부작용)와 건강 편익(예: 삶의 질 향상)의 균형을 어떻게 고려하였는지 기술한다.	

부록 1.

NECA
이해관계
선언(COI)
양식

부록 2.

체계적
문헌고찰
질
평가도구
: AMSTAR

부록 3.

무작위
배정연구
질
평가도구
Cochrane
ROB

부록 4.

비무작위
연구의
질
평가도구
RoBANS

부록 5.

임상진료
지침 표준
보고가이드
도구
STARIGs

구분/항목	번호	검토 항목	보고된 페이지
권고안 도출방법	20	권고안 도출에 사용한 방법(예: 명목집단 기법)을 기술한다.	
동료검토 절차	21	진료지침 출간 이전에 시행한 동료검토 절차를 기술하고, 가능한 경우 동료검토 그룹의 인적사항(예: 이름, 소속기관명, 직위)을 제시한다.	
권고와 해당근거			
권고안	22	대상 인구집단에서 특정 상황에 권장되는 행위를 제시한다.	
대체 가능한 치료 방법	23	진료지침이 다루는 특정 증재의 선택 가능한 대안을 권고안 혹은 해당근거에 제시한다.	
동료검토	24	동료검토를 통해 얻어진 결과/정보는 무엇이고, 진료지침 권고안 결정에 미친 영향을 기술한다.	
근거의 제한점	25	권고안 도출에 사용된 근거의 제한점(예: 근거 부족, 근거의 이질성)을 기술한다.	
연구개발 필요 주제 제시	26	권고안 도출시 근거가 부족한 영역, 향후 연구가 필요한 주제를 기술한다.	
독립성			
개발기금	27	개발기금 출처(예: 보건복지부, 학회자체 예산)와 개발과정에 미친 영향을 기술한다.	
이해관계 선언	28	개발그룹 구성원들의 진료지침과 관련된 이해관계를 다룬 방법과 처리한 결과를 기술한다.	
실행 및 확산			
알고리즘	29	단계별 의사결정이 필요한 경우 알고리즘을 제시한다.	
확산도구	30	진료지침 확산도구(예: 환자용 지침, 간이 진료지침, 요약본)를 부록문서로 제시한다.	
갱신 계획			
갱신 계획	31	진료지침 갱신 계획 유무와 관련 계획(시점, 기준, 방법 등)을 기술한다.	
관련 정보			
용어정의	32	진료지침의 올바른 적용을 위하여 해설이 필요한 용어가 있는 경우, 이에 대한 정의와 약어에 대한 정보를 기술한다.	
참고문헌	33	진료지침 개발에 사용된 참고문헌을 기술한다.	
부록	34	덧붙임이 필요한 정보(예: 세부적인 검색전략, 이해관계 선언 내용, 핵심질문)를 기술한다.	



NECA 한국보건의료연구원
National Evidence-based Healthcare Collaborating Agency



9 788994 581644 9 3 5 1 0
ISBN 978-89-94581-64-4