

리빙 가이드라인과 기관지 확장증 진료지침의 개발 방향

방법론적 엄격성과 효율성의 설계

한국코크란/고려대학교 의과대학
김현정

강의 구성

I

리빙 구조

거버넌스, 갱신 트리거
, 개발기록

II

방법론적 엄격성

사전계획, 근거평가,
판단의 구조화

III

효율성

통합검색, 모듈형 갱신
, 규칙화 범위

IV

운영과 지속

근거감시, 버전관리,
실행평가

- 다루는 범위: 지침 개발방법론
- 다루지 않는 범위: 기관지확장증 개별 권고의 임상적 내용, 약제 선택의 임상적 판단

리빙 가이드라인의 정의

전통적 개정 모델

- 3-5년 주기의 전면 개정
- 개정 시점의 개발팀 재구성
- 발간 시점과 근거 시점 간 시차 발생
- 개정 사유와 변경 내역의 기록 부재
- 유효 버전 식별의 어려움

리빙 모델

- 영향받는 권고 단위의 선택적 재평가
- 근거감시-합성-판단의 상시 기능화
- 권고 단위의 버전 및 변경 사유 명시
- 최초 개발 규칙의 반복 적용
- 현재 유효 버전의 상시 식별

- 갱신 단위는 지침 전체가 아니라 개별 권고
 - 리빙의 요건은 검색 빈도가 아니라 권고를 재평가할 수 있는 절차의 존재
 - Akl 2017, J Clin Epidemiol — Living systematic reviews 4: living guideline recommendations

대상 영역의 선정: 기관지확장증

질병부담

- 반복 악화와 입원
- 장기간 지속되는 삶의 질 저하

근거의 축적

- 장기 macrolide, 흡입 항생제
- 기도청결요법, 신규 약물계열
- 관련 임상시험의 지속적 발표

국내 진료환경

- 급여기준과 접근성의 차이
- 국제 지침의 직접 적용 제약

- 리빙 적용의 전제조건
 - 우선순위가 높은 질문 / 근거가 불확실한 질문 / 신규 근거가 계속 유입되는 질문
 - 세 조건은 지침 전체가 아니라 핵심질문 단위로 판정

핵심질문의 트리아지: 필요성과 기준

- 리빙 운영은 지침 전체가 아니라 핵심질문 단위로 적용
 - 질문별 분류가 감시 범위와 갱신 자원의 배분 기준이 됨
- 분류가 필요한 이유
 - 자원의 유한성 — 모든 질문을 동일 빈도로 상시 감시하는 것은 실행 불가
 - 근거 유입 속도의 차이 — 질문별로 신규 임상시험의 발생률이 다름
 - 갱신 산출의 차이 — 권고 변경 가능성이 낮은 질문의 반복 재검토는 산출이 낮음
 - 트리거 정의의 전제 — 감시 소스, 확인 빈도, 재검토 조건이 분류에 따라 달라짐

분류 판정에 사용하는 네 가지 요소

① 현재 근거의 확신도 ② 진행 중 연구의 존재와 규모 ③ 권고 변경 시 진료에 미치는 영향 ④ 임상적 논쟁의 정도

기관지확장증 핵심질문의 분류

● Living

상시 감시 대상

KQ3 장기 흡입 항생제

KQ6 녹농균 박멸요법

KQ9 DPP-1 억제제

- 현재 확신도가 낮음
- 다수의 임상시험 진행 중
- 권고 변경 시 진료 영향이 큼

● Periodic

연 1회 정기 검색 대상

KQ1 기도청결요법

KQ2 점액용해제

KQ4 장기 macrolide

KQ8 폐재활

- 근거 기반이 어느 정도 확립
- 신규 연구의 발생률이 낮음
- 세부 조건 변경 수준의 갱신 예상

● Static

차기 전면 개정까지 유지

KQ5 장기 경구 항생제

KQ7 장기 흡입 스테로이드

KQ10 객혈에 대한 색전술

- 근거 지형의 변화가 관찰되지 않음
- 대규모 신규 시험 계획이 확인되지 않음
- 무작위 연구 수행이 어려운 영역 포함

- 본 분류는 근거 지형에 따른 예시이며, 확정 = 개발위원회의 판정 사항
- 분류는 고정값이 아니라 갱신 시마다 재판정 — Periodic에서 대규모 시험 발표 시 Living으로 상향

분류의 가변성: 두 가지 사례

KQ4 장기 macrolide

현재 Periodic

분류 근거 효과에 관한 근거 기반이 비교적 확립되어 있음

재판정 요인 장기 사용에 따른 내성 등 안전성 자료의 축적

트리거의 성격 효과추정치의 변화가 아니라 안전성 신호의 발생

이동 방향 안전성 신호 확인 시 Periodic → Living

KQ10 객혈에 대한 색전술

현재 Static

분류 근거 근거가 안정적이어서가 아니라 무작위 연구가 생성되기 어려운 구조

재판정 요인 등록자료의 축적, 시술 기법과 기기의 변화

트리거의 성격 포함 연구설계 범위의 재검토(관찰연구 포함 여부)

이동 방향 자료원 변화 시 Static → Periodic

■ 분류 가변성의 의미

- Static은 단일 범주가 아님 — 근거가 안정된 경우와 근거가 생성되기 어려운 경우를 구분
- 분류 판정은 단일 축의 작업이 아님 — 신규 근거 확인은 방법론 축, 임상적 의미 판단은 임상 축
- 분류가 갱신 시마다 달라지므로 판정을 수행할 상시 기능이 필요 — 리빙 구조를 두는 실질적 이유

개발 원칙과 강의의 두 축

투명성

사전계획

재현가능성

방법론적 엄격성

효율성

방법론적 엄격성

- 정의: 동일한 자료와 규칙에서 동일한 결론이 산출되는 성질
- 구현: 독립 이중평가, RoB 2/ROBINS-I, GRADE, EtD
- 리빙에서의 추가 조건: 갱신 시점마다 재현

효율성

- 정의: 중복작업과 불필요한 변이의 감소
- 구현: 통합검색, 역할분담, 규칙의 사전 구조화
- 위치: 엄격성의 대립항이 아니라 지속조건

리빙 구조

01

거버넌스 · 갱신 트리거 · 개발기록

개발 절차와 갱신의 관계

갱신은 새로운 절차가 아니라 최초 개발 절차의 부분 재실행에 해당

핵심질문 > 결과 선정 > 근거합성 > GRADE > EtD > 권고투표 > 갱신

각 단계의 판단 규칙이 기록된 경우

→ 갱신 시 동일 규칙을 재적용

기록되지 않은 경우

→ 해당 단계의 판단을 재수행

결과

→ 갱신 비용은 최초 개발 시점의 문서화 수준에 의해 결정

거버넌스 구성

Clinical stewardship

- 지침의 범위와 핵심질문 설정
- 중요한 결과의 임상적 해석
- EtD 토론, 권고의 방향·강도 결정
- 신규 근거의 임상적 재평가 필요성 판단
- 확산, 실행, 장기 유지

Evidence & methods stewardship

- 검색전략 수립과 문헌선택
- 비뚤림 위험 평가
- 근거합성 및 GRADE 평가
- Evidence profile, SoF 작성
- 갱신 시 동일 기준·방법의 적용

- 두 축의 관계
 - 근거가 변화하면 방법론 축에서 근거를 갱신하고, 권고 재평가가 필요하면 임상 축이 다시 활성화
 - 책임의 분리이며 권한의 이전은 아님 — 임상적 판단은 방법론 조직으로 이전되지 않음

개발 주체의 지속성

- 1 연구사업의 한시성 과제 종료와 함께 개발위원회가 해체되는 경우가 많음
 - 2 인적 구성의 변동 갱신 시점에 최초 개발 참여자가 남아 있다고 가정하기 어려움
 - 3 방법론적 연속성 판단 규칙을 조직이 보유할 때 위원 구성이 바뀌어도 절차가 재현됨
 - 4 상시 역량 리빙의 요건은 갱신 빈도가 아니라 적시 재평가가 가능한 조직 역량
- 지속가능성의 설계 지점은 개별 위원의 참여 지속이 아니라 기능의 조직 귀속

갱신 트리거의 정의

신규 근거
확인

>

효과추정치·확신
도
변화 여부

>

권고 재평가
필요 여부

>

EtD 재검토

>

권고 유지 /
변경

방법론 축의 판단 대상

- 신규 연구가 기존 효과추정치를 변화시키는가
- GRADE 확신도 등급이 달라지는가

임상 축의 판단 대상

- 변화한 근거가 권고의 방향을 바꾸는가
- 권고의 강도를 조정할 필요가 있는가

- 트리거는 사전에 정의된 조건을 의미하며, 신규 문헌의 출현 자체는 트리거에 해당하지 않음
- 명시가 필요한 항목: 감시 소스, 담당 역할, 확인 빈도, 재검토 개시 조건

개발기록과 추적성

검색기록

검색식, 실행일자, 검색 결과 수

문헌선택

2인 독립 선택 결과, 불일치
해소 경위

비뚤림 평가

영역별 판정과 판정 근거 기술

Evidence profile · SoF

효과추정치, 확신도와 판단 사유

EtD 판단

항목별 위원회 판단과 근거자료

투표결과

전자투표 원자료, 합의 판정 결과

- 관리 방식: 공동 전자저장소와 지침 개발플랫폼
- 전자투표의 목적: 원자료와 최종 기록 사이의 수기 전사 제거 — 오류 감소 및 갱신 시 재사용

방법론적 엄격성

02

사전계획 · 근거평가 · 판단의 구조화

방법론적 엄격성의 조작적 정의

정의

동일한 자료와 동일한 규칙이 적용될 때 동일한 결론이 산출되는 성질

리빙에서의 조건 수행 시점과 수행자가 달라져도 동일한 결론이 재현되어야 함

귀결 개인의 숙련도가 아니라 절차 설계의 속성으로 정의됨

구현 수단 판단 규칙의 사전 정의, 규칙 적용 과정의 기록

사전계획 ① 중요한 건강결과의 선정

근거합성 결과를 확인한 이후 결과의 중요도가 변경되는 것을 최소화하기 위해 근거평가 전에 결정한다.

7-9점

의사결정에 결정적인 결과

4-6점

중요하지만 결정적이지 않은 결과

1-3점

덜 중요한 결과

절차

- 근거평가 착수 전에 시행
- 결과별 중요도를 1-9점으로 평정
- 결정적 결과와 중요한 결과로 분류
- SoF와 EtD의 대상 결과를 고정
- 갱신 시에도 동일한 결과 목록 사용

- GRADE guidelines 2 — Framing the question and deciding on important outcomes (Guyatt 2011)

기관지확장증에서의 결과 후보

악화 빈도

결정적 결과 후보

악화로 인한 입원

결정적 결과 후보

질병특이 삶의 질

QOL-B, SGRQ 등

폐기능 (FEV₁)

중요한 결과

객담량 · 미생물학적 반응

중간지표 — 해석에 주의

이상반응 · 내성 발생

중요한 결과

사전계획 ② 임상적 중요성 기준

설정 목적: 통계적으로 유의한 효과와 임상적으로 중요한 효과를 구분하기 위한 사전 기준 확보

최대 3

합의조사 라운드

$\geq 70\%$

응답률

$> 60\%$

동의율

$< 20\%$

비동의율

확정된 threshold의 사용처

- 효과크기의 해석 기준
 - GRADE 비정밀성 판정의 경계값
 - EtD의 desirable / undesirable effects 판단 기준
-
- 사전 확정 시 갱신 시점에 기준을 재논의할 필요가 없어, 사전계획은 효율성과 직접 연결됨

MCID와 지침 threshold의 구분

MCID / MID

- 측정 수준: 개별 환자
- 산출 맥락: 특정 집단, 특정 척도, 특정 측정방식
- 제약: 직접 적용 가능한 값이 항상 존재하지는 않음

지침 수준 threshold

- 측정 수준: 집단 수준 의사결정
- 설정 방식: 문헌 근거와 전문가 합의
- 고려 요소: 대상 환자, 결과척도, 효과측정 방식

- 본 지침에서는 두 값이 항상 동일하다는 전제를 두지 않음
 - MCID/MID 자료는 threshold 설정의 정보원으로 사용하되, 지침 수준의 판단은 별도로 수행
 - 해당 기준은 절대적 biological threshold가 아니라 의사결정을 위한 prespecified threshold로 해석

임상적 중요성 역치를 설정하는 이유

출발점: 역치 설정의 제약

- 다수의 결과에서 신뢰할 만한 MCID / MID가 존재하지 않음
- MCID / MID는 개인 내 변화의 최소 중요 차이이며, 두 군의 효과 차이를 판정하기 위한 값이 아님

1 결정과정의 투명성

- 모든 치료는 이득과 위해를 함께 가짐
- 어느 정도의 이득을 의미 있다고 보았는지가 권고의 실질적 근거
- 역치가 없으면 권고는 남고 판단 기준은 남지 않음
- 역치의 명시는 곧 판단 기준의 공개

2 임상현장에서의 재적용

- 환자와 임상가는 각자의 역치를 가짐
- 위원회 역치가 공개되면 자신의 기준과 대조 가능
- 역치가 다르면 동일한 근거에서도 다른 선택이 가능
- 권고 이행 여부를 스스로 재결정할 수 있음

- 역치는 효과의 참값을 규정하는 값이 아니라, 위원회 판단의 기준선을 드러내는 장치
- 따라서 절대적 biological threshold가 아니라 의사결정을 위한 prespecified threshold로 해석

근거검색과 문헌선택

PICO 확정
(임상전문가)



검색식 개발
(검색전문가)



예비검색 후
전략 조정



제목·초록
2인 독립



원문검토
2인 독립



PRISMA
보고

검색원

MEDLINE, Embase, Cochrane Library,
KoreaMed, ClinicalTrials.gov

불일치의 처리

논의로 해결하며, 합의되지 않으면 제3의 평가
자 또는 방법론 전문가가 검토

- 검색식은 갱신 시 동일한 형태로 재실행되므로 재현 가능한 형태로 기록
- 연구설계 제한: 근거가 충분한 질문은 RCT로 한정, 제한적인 질문은 설계를 제한하지 않음

비뚤림 위험 평가

RoB 2

- 대상: 무작위배정 비교임상시험
- 단위: 가능한 경우 결과별 평가
- 방식: signalling question 기반 알고리즘 판정

ROBINS-I

- 대상: 비무작위 중재연구
- 기준: 목표시험(target trial) 대비 평가
- 요건: 교란변수와 선택비뚤림의 사전 명시

- 평가자 간 변이의 관리
 - 각 핵심질문별 2인이 독립 평가한 후 대조하고, 불일치는 합의로 조정
 - 평가 이전에 영역 정의, signalling question 해석, 판정원칙, 근거기술 방식에 대한 공통교육 시행
 - 갱신 시 평가자가 변경될 수 있으므로 교육자료를 조직 자산으로 보관

근거 확신도의 평가: GRADE

하향요인

비뚤림 위험

비일관성

비직접성

비정밀성

출판비뚤림

비정밀성 판정의 규칙화

- 95% 신뢰구간이 사전 설정한 임상적 중요성 경계를 교차하는지 확인
- Optimal information size(OIS) 충족 여부를 함께 고려
- 두 조건을 판정 알고리즘으로 사전에 표현하면 동일 입력에 동일 판정이 반복됨

상향요인 · 큰 효과 · 용량-반응 관계 · 잔여교란의 방향

- 규칙으로 표현 가능한 판단은 규칙화하고, 표현이 어려운 판단에 전문가 시간을 배분
- 이 구분이 이후 효율성 파트의 전제가 됨

환자의 가치와 선호도

GRADE에서 가치와 선호도는 주요 건강결과에 환자가 부여하는 상대적 중요성을 의미하며, 권고의 방향과 강도에 영향을 준다.

조사 내용

- 질병으로 인한 어려움
- 치료에 기대하는 효과
- 치료결정 시 고려 요소
- 지불의향, 치료성공의 기준

조사 대상

- 환자 대상 다기관 단면조사
- 의료진 대상 별도 조사 병행
- 두 자료를 구분하여 제시

EtD에서의 사용

- values and preferences 판단의 직접 근거
- 효과근거와 가치판단을 분리하여 제시

- 위원회의 추정치를 사용하는 대신 국내 환자 자료를 직접 생산하여 투입
- GRADE guidelines 19 (Zhang 2019)

EtD 판단과 권고투표의 분리

EtD criterion별
판단

권고 방향
투표

합의 시
강도 투표

미합의 시
다음 회의로 이월

투표 성립요건

전체 위원의 70% 이상 참석

합의기준

권고의 방향과 강도 각각 투표 참여 위원의 70% 이상 동의

EtD criterion

중간 판단이므로 다수 판단을 사용, 동률인 경우에만 분포 공개 후 추가 논의

투표방식

다른 위원의 선택을 확인할 수 없는 익명 독립투표

진행자

지침 내용에 이해관계가 없는 방법론 전문가, 임상적 의견은 제시하지 않음

- 미합의 질문은 동일 회의에서 재투표하지 않고 다음 회의로 이월
- 합의의 기능: 불확실성의 제거가 아니라 불확실성 하에서 내린 판단의 명시적 기록

효율성

03

중복작업과 번이의 감소

효율성의 조작적 정의

정의

방법론적 엄격성을 저하시키지 않으면서 중복작업과 불필요한 변이를 줄이는 것

대상

판단 자체가 아니라 판단에 이르는 반복작업과 절차상의 변이

수단

반복 가능한 작업의 통합과 표준화, 판단 규칙의 사전 구조화

귀결

전문가의 시간을 판단이 필요한 단계에 재배분 — 엄격성의 지속조건

통합검색 전략

동일한 인구집단과 유사한 중재를 다루는 핵심질문들은 질문별로 검색할 경우 상당한 중복 발생

질문별 독립 검색

KQ 1

검색 · 선별 반복

KQ 2

검색 · 선별 반복

KQ 3

검색 · 선별 반복

통합검색 후 분기

공통 문헌 pool
중복제거

KQ 1 분기

KQ 2 분기

KQ 3 분기

- 효과: 검색과 선별의 반복 감소, 둘 이상의 질문에 관련되는 연구의 누락 가능성 감소
- 리빙 가이드라인 검색방법에서 제시된 single-search approach (McDonald 2023)

모듈형 갱신

신규 근거의 영향이 특정 핵심질문에 국한되는 경우, 전체 지침을 다시 평가하지 않고 해당 질문 만 선택적 재검토



- 재검토 범위: 해당 질문의 근거합성 → 확신도 평가 → EtD → 권고
- 예외: 복수 질문에 동시에 영향을 주는 근거(공통 안전성 신호 등)는 별도의 처리 규칙 필요

작업의 규칙화 범위

규칙화 가능한 작업

- 검색식의 재실행과 중복제거
- 선별 우선순위의 지정
- 자료추출 초안의 생성
- 사전 규칙에 따른 비정밀성 판정
- 버전 간 변경내역의 생성

판단이 유지되어야 하는 영역

- 결과의 임상적 중요성
- 비직접성의 해석
- 이득과 위해의 균형
- 권고의 방향과 강도
- 갱신 필요성에 대한 최종 결정

- 자동화의 기능: 판단의 대체가 아니라 판단이 필요한 지점의 조기 식별
- 개발 지원도구의 정확성과 효율성 자체는 지침 개발과 분리된 별도 연구로 검증

근거감시 체계

정기 검색

매년 정해진 시점에 검색전략을 동일한 형태로 재실행

수시 재검토

권고를 변경할 가능성이 있는 근거가 확인되면 해당 질문을 즉시 재검토

정기 검색 사이의 감시를 규정하는 네 가지 항목

- 감시 소스 — 대상 데이터베이스, 알림 서비스, 학술대회 발표 등
 - 담당 주체 — 개인이 아니라 역할 단위로 지정
 - 감시 빈도 — 확인 주기
 - 재검토 개시 조건 — 권고를 다시 여는 기준
- 정기적 개정과 리빙을 구분하는 실질적 요소는 이 네 항목의 명시 여부

출판 경로와 버전 관리

학술지 출판본 — archival version

- 특정 시점에 승인된 지침을 고정
- 동료심사, 인용, 영구 기록의 기능
- 발간 이후 수정되지 않음

웹 기반 — current living version

- 현재 유효한 권고를 제공
- 버전 공개시점, 변경 질문, 변경 사유 표시
- 이전 버전과의 차이 확인 가능

- 두 경로는 대체 관계가 아니라 보완 관계
 - 이용자 관점의 요건: 현재 유효한 권고를 식별할 수 있을 것
 - 사전 정의가 필요한 사항: 버전 번호 부여 규칙(권고 변경 없음 / 강도 변경 / 방향 변경)

실행 및 영향평가

지침 발간 이후의 실제 활용을 확인하기 위해 평가지표를 발간 전에 정의한다.

과정지표

적응 대상 환자 중 해당 치료
가 실제로 고려된 환자의 비율

결과지표

치료 후 의미 있는 악화 감소
등 임상적 호전 여부

도달지표

지침의 인지도 및 활용 정도,
즉 지침 도달률

- 사전 정의의 이유: 발간 후에 지표를 설정하면 비교 기준시점을 확보할 수 없음
- 평가 범위: 출판 여부가 아니라 임상실천과 건강결과에 미치는 영향

리빙 구조를 유지하는 운영 인프라

Notion

의사소통과 자료 저장

- 위원 간 논의와 일정의 단일 창구
- 회의자료, 초안, 버전 이력의 추적
- 참여자가 바뀌어도 맥락이 남음

한국코크란

지침 개발 플랫폼

- 메타분석, 근거의 확신도 관리
- 핵심질문 단위의 근거표 보관
- 갱신 시 이전 판단의 재조회

전용 앱

통합 투표 시스템

- 중요 건강결과 평정
- 내부위원회 EtD 및 권고 투표
- 외부검토, 공청회 투표

세 시스템의 연계

투표 시스템에서
의사결정 수행

결과가 URL을 통해
자동 전달

가이드라인 플랫폼에
결과 적재

발표자료 · 문서
자동 생성

- 연계의 목적은 수기 전사와 재입력의 제거
 - 투표 원자료와 최종 기록 사이의 불일치 가능성이 발생하지 않음
 - 동일한 자료가 근거표, 회의자료, 원고 초안으로 반복 사용됨
 - 효율성은 절차의 축소가 아니라 반복 작업의 제거에서 확보됨

요약

리빙 가이드라인은 근거가 변화할 때 재검토 대상을 식별하고,
동일한 방법론적 구조 안에서 권고를 다시 결정할 수 있는 개발체계 의미

- 1 판단 기준의 사전 정의
- 2 판단 근거와 수정 이력의 기록
- 3 근거생산과 임상판단의 상시 주체 확보
- 4 변경된 부분의 선택적 재평가
- 5 현재 유효 버전의 상시 식별