NECA-의료기술재평가사업
NECA-R-25-001-40

의료기술재평가보고서 2025

덴버발달검사

NECA
한국보건의료연구원

의료기술재평가사업 총괄

김민정 한국보건의료연구원 보건의료평가연구본부 본부장

전미혜 한국보건의료연구원 보건의료평가연구본부 재평가사업팀 팀장

연 구 진

담당연구원

정자혜 한국보건의료연구원 재평가사업팀 연구원

부담당연구원

박은정 한국보건의료연구원 재평가사업팀 연구위원

주 의

1. 이 보고서는 한국보건의료연구원에서 수행한 의료기술재평가사업(NECA-R-25-001)의 결과보고서입니다.

2. 이 보고서 내용을 신문, 방송, 참고문헌, 세미나 등에 인용할 때에는 반드시 한국보건의료연구원에서 수행한 평가사업의 결과임을 밝혀야 하며, 평가내용 중 문의사항이 있을 경우에는 주관부서에 문의하여 주시기 바랍니다.

차 례

요약문(국문) ····· i
알기 쉬운 의료기술재평가 ····· iv

Ⅰ. 서론 ····· 1
1. 평가배경 ····· 1
1.1 평가대상 의료기술 개요 ····· 1
1.2 국내외 보험 및 행위등재 현황 ····· 5
1.3 국내 이용 현황 ····· 7
1.4 의학교과서 ····· 7
1.5 관련 국내외 가이드라인 ····· 7
1.6 관련 체계적 문헌고찰 ····· 8
1.7 관련 의료기술 ····· 9
2. 평가목적 ····· 16

Ⅱ. 평가방법 ····· 17
1. 문헌고찰 ····· 17
1.1 개요 ····· 17
1.2 핵심질문 ····· 17
1.3 연구검색 ····· 17
1.4 연구선정 ····· 18
1.5 자료정리 및 분석 ····· 19
2. 전문가 의견 ····· 19
2.1 개요 ····· 19
2.2 전문가 추가 자문 ····· 19
2.3 검토 내용 ····· 19
2.4 의견 수렴 및 반영 방법 ····· 20
3. 결론 도출 ····· 20

Ⅲ. 평가결과 ····· 21
1. 문헌고찰 ····· 21
1.1 연구선정 개요 ····· 21
1.2 선택연구 특성 ····· 22
1.3 검사도구의 타당성 ····· 23
1.4 선별검사로서의 정확도 ····· 30
2. 전문가 의견 ····· 35

Ⅳ. 결과요약 및 결론 ····· 36
1. 평가결과 요약 ····· 36
2. 결론 및 제언 ····· 37

Ⅴ. 참고문헌 ····· 39

Ⅵ. 부록 ····· 41
1. 의료기술재평가위원회 ····· 41
2. 소위원회 ····· 42
3. 연구검색 현황 ····· 43
4. 최종선택연구 ····· 45

표 차례

표 1.1 발달지연의 정의 ····· 2
표 1.2 영유아건강검진 발달평가 도구 관련 보건복지부 고시 이력 ····· 3
표 1.3 건강보험 요양급여·비급여 비용 목록 등재 현황(덴버발달검사) ····· 5
표 1.4 건강보험심사평가원 고시항목 상세 ····· 6
표 1.5 국외 보험 및 행위 등재 현황 ····· 6
표 1.6 관련 체계적 문헌고찰 ····· 8
표 1.7 영유아 발달평가 검사의 활용 예시 ····· 10
표 1.8 건강보험 요양급여·비급여 비용 목록 등재 현황(베일리영아발달측정) ····· 12
표 1.9 건강보험심사평가원 행위기술서(베일리영아발달측정) ····· 12
표 1.10 베일리영아발달측정 관련 보건복지부 고시 이력 ····· 13
표 1.11 상대가치점수 ····· 14
표 1.12 베일리영아발달측정 이용 현황 ····· 14
표 1.13 영유아 발달검사의 목적 및 검사 특성 비교 ····· 15
표 1.14 발달검사의 연령별 적용 가능 구간 ····· 16
표 2.1 국외 전자데이터베이스 ····· 17
표 2.2 국내 전자데이터베이스 ····· 18
표 2.3 연구의 선택 및 배제기준 ····· 18
표 3.1 선택연구의 특성 ····· 22
표 3.2 덴버발달검사의 타당성 ····· 26
표 3.3 덴버발달검사의 선별검사로서의 정확도 ····· 32

그림 차례

그림 1.1 베일리영아발달측정 환자 수 추이 ····· 13
그림 3.1 연구선정 흐름도 ····· 21

요약문(국문)

평가 배경

덴버발달검사(Denver Developmental Screening Test II, 이하 Denver II)는 생후 1개월부터 만 6세까지의 영유아를 대상으로 대근육운동, 미세운동-적응, 언어, 개인-사회성 등 4개 발달 영역을 평가하여 발달지연 아동을 조기에 발견하고 정밀검사 또는 중재로 연계하기 위한 선별용 발달검사이다. 동 의료기술은 신의료기술평가 제도 도입 이전인 2001년 비급여 항목으로 등재되어 현재 임상에서 사용되고 있다. 이후 동 검사의 사용 현황과 유용성에 대한 근거를 검토할 필요성이 제기되었으며, 2025년 제7차 의료기술재평가위원회(2025. 7. 11.)에서 재평가 계획서 및 소위원회 구성안을 심의한 후 재평가를 수행하였다.

평가 목적

본 평가의 목적은 발달지연이 의심되는 영유아에서 덴버발달검사의 임상적 유용성 등에 대한 근거를 종합적으로 검토하여, 향후 관련 의사결정에 활용할 수 있는 근거를 제공하고자 수행되었다.

평가 방법

본 평가는 “덴버발달검사는 발달지연이 의심되는 영유아를 선별하는 데 임상적으로 유용성이 있는가?”라는 핵심질문을 기반으로 문헌고찰과 전문가 자문을 병행하여 수행하였다. 문헌검색은 국외 2개(Ovid MEDLINE, APA PsycInfo (via Ovid))와 국내 1개(KoreaMed) 데이터베이스에서 수행하였으며, 수집된 문헌은 질적 검토(Qualitative review)를 통해 정리하였다. 또한 문헌고찰만으로 확인하기 어려운 국내 임상현장의 실제 활용 현황, 검사 결과의 활용 방식, 기존 발달선별검사와의 관계 등을 보완적으로 확인하기 위하여 전문가 자문을 실시하였다. 모든 평가방법은 평가목적을 고려하여 “덴버발달검사 소위원회(이하 ‘소위원회’라 한다)”의 심의를 거쳐 확정하였다. 소위원회는 재활의학과(소아) 1인, 소아청소년과 2인, 정신건강의학과(소아) 1인, 임상심리전문가 2인, 근거기반의학 1인의 총 7인으로 구성하였다.

평가 결과

문헌고찰 결과, 핵심질문을 충족하는 총 15편의 문헌이 선택되었다. 선택문헌은 평가목적에 따라 검사도구로서의 타당도 및 신뢰도를 보고한 연구 8편과 선별검사로서의 정확도를 보고한 연구 8편으로 구분되었으며, 이 중 1편은 두 범주에 중복 포함되었다. 또한 문헌고찰만으로 확인하기 어려운 국내 활용 맥락, 국외 활용 동향, 한국 영유아 발달선별검사(Korean Developmental Screening Test for Infants and Children, K-DST) 도입 이후의 역할 등을 보완적으로 검토하기 위하여 전문가 자문 결과를 함께 검토하였다.

검사도구의 타당성

검사도구로서의 타당도 및 신뢰도 연구에서는 덴버발달검사의 개발 및 재표준화, 국가별·문화권별 규준 비교, 번역·적응판의 신뢰도 및 타당도 검증 결과가 보고되었다. 검토 결과, 덴버발달검사는 문항 구성과 연령별 통과연령 산출 과정을 거쳐 개발·표준화되었으며, 여러 연구에서 검사자 간 신뢰도와 검사-재검사 신뢰도는 전반적으로 양호한 수준으로 보고되었다. 다만 국가별·문화권별 표준화 연구에서는 일부 문항의 통과연령과 규준 차이가 확인되어, 문화적·언어적 맥락을 고려한 해석과 지역별 표준화의 필요성이 제시되었다.

선별검사로서의 정확도

선별검사로서의 정확도 연구에서는 일반 아동, 발달지연 의심 또는 위험 아동, 미숙아 등 고위험군을 대상으로 덴버발달검사의 민감도, 특이도, 예측도, 검사 간 일치도 및 선별률 등이 보고되었다. 덴버발달검사는 여러 연구에서 발달지연 위험군을 조기에 확인하는 데 활용 가능한 선별도구로 보고되었으나, 연구대상자, 참고표준 및 판정 기준에 따라 정확도 지표에는 차이가 있었다. 일부 연구에서는 비교적 높은 민감도가 보고된 반면 특이도는 낮거나 변동이 있어, 검사 결과만으로 발달지연 여부를 확정하기보다는 정밀 발달평가 필요성, 추적관찰 여부 및 임상적 판단을 함께 고려할 필요가 있는 것으로 제시되었다.

전문가 의견

본 평가에서는 문헌고찰 결과를 보완하고 국내 임상현장에서의 활용 현황 및 보건의료체계 내 적용 맥락을 확인하기 위하여 관련 분야 전문가를 대상으로 추가 자문을 실시하였다. 전문가 자문에서는 덴버발달검사의 검사 성격, 국내외 활용 동향, 지역사회 기반 활용 사례, K-DST 도입 이후의 역할 등이 검토되었다. 자문 결과, 덴버발달검사는 특정 발달 문제의 심각성을 판단하거나 이상 여부를 진단하기 위한 도구가 아니라, 아동 발달의 전반적 흐름을 확인하고 발달지연 가능성이 있는 아동을 선별하기 위한 도구로 보는 것이 적절하다는 의견이 제시되었다.

또한 덴버발달검사는 검사자가 아동의 행동과 발달 상태를 직접 관찰하면서 시행할 수 있어, 일부 가정방문사업 등 지역사회 기반 사업에서도 활용된 사례가 확인되었다. K-DST 도입 이후에는 보호자 보고 기반 선별검사와 관찰 기반 선별검사의 역할 차이를 고려하여, 임상 및 지역사회 상황에 따라 병행 또는 선택적으로 활용될 경우 선별정확도를 보완하는 데 도움이 될 수 있다는 의견이 제시되었다.

결론 및 제언

본 의료기술재평가는 덴버발달검사의 국내 임상적 유용성을 평가하기 위하여 문헌고찰과 전문가 의견을 종합적으로 검토하였다. 문헌고찰 결과, 덴버발달검사는 발달지연 위험군을 조기에 선별하기 위한 도구로 활용될 수 있는 제한적 근거가 확인되었다. 다만 선별정확도와 임상적 유용성에 대한 국내 근거는 연구 수, 연구 규모 및 대상 집단 측면에서 제한적이었으며, 다양한 임상 환경을 충분히 대표한다고 보기는 어려웠다.

소위원회는 덴버발달검사가 장애 진단이나 발달 상태를 확진하기 위한 진단도구가 아니라, 발달지연이 의심되는 아동을 대상으로 조기 선별 및 전반적인 발달 상태를 확인하기 위한 선별검사라고 판단하였다. 또한 검사 결과는 단독으로 해석하기보다 정밀평가 필요성, 추적관찰 여부, 임상적 판단 등을 종합적으로 고려하기 위한 참고자료로 활용하는 것이 적절하다고 보았다. 실제 활용 양상은 진료과 및 의료기관 유형, 진료 단계, 의료자원 여건, 대상 아동의 특성에 따라 달라질 수 있으며, K-DST 등 기존 발달선별검사와의 관계에서도 보호자 보고 기반 선별 결과를 임상가의 직접 관찰로 보완하는 역할을 할 수 있다고 보았다. 아울러 동 검사의 임상적 위치와 활용 범위를 보다 명확히 평가하기 위해서는 실제 보건의료 현장에서의 사용 현황과 검사 결과의 활용 과정을 체계적으로 파악할 수 있는 자료의 축적·수집·분석이 필요하다고 판단하였다.

2026년 제2차 재평가전문위원회*(2026. 2. 6.)는 '덴버발달검사'에 대해 다음과 같이 심의하였다. 재평가전문위원회는 소위원회의 검토 결과에 동의하였으며, 덴버발달검사가 스크리닝(screening) 또는 지속적 추적관찰(surveillance) 차원에서 최소한의 발달 이상 여부를 확인하는 검사로서 갖는 의미와 지속 필요성에 대해서는 향후 추가 근거 검토를 통한 재평가가 필요하다는 의견을 제시하였다.

*「신의료기술평가에 관한 규칙」(보건복지부령 제1098호, 일부개정, 2025.9.7. 시행) 개정으로 재평가전문위원회가 새로 구성되어 2025년 9월부터 운영되고 있다.

주요어

덴버발달검사, 발달선별검사, 발달지연, 임상적 유용성

Denver Developmental Screening Test II (Denver II), Developmental Screening Test, Developmental Delay, Clinical Utility

알기 쉬운 의료기술재평가

덴버발달검사는 발달지연이 의심되는 영유아를 선별하는 데 임상적으로 유용한가요?

질환 및 의료기술

덴버발달검사(Denver II)는 생후 1개월부터 만 6세까지의 영유아가 또래에 비해 적절하게 발달하고 있는지를 확인하는 검사이다. 운동, 언어, 사회성 등 4개 영역의 발달 상태를 평가하여 발달지연 가능성을 선별하는 데 사용되며, 현재 비급여로 사용되고 있다.

의료기술의 임상적 유용성

덴버발달검사의 임상적 유용성은 관련 연구 15편과 전문가 자문 결과를 종합하여 검토하였다.

검토 결과, 덴버발달검사는 여러 국가에서 사용되고 있으며 검사 결과의 일관성도 전반적으로 양호한 것으로 확인되었다. 또한 일반 영유아뿐 아니라 발달지연이 의심되는 아동이나 미숙아 등에서 발달 상태를 확인하고 추가 평가가 필요한 아동을 선별하는 데 활용될 수 있는 것으로 보고되었다.

다만 검사 결과만으로 발달장애를 확정적으로 진단할 수는 없었으며, 선별정확도는 연구에 따라 차이가 있었다. 따라서 덴버발달검사는 발달장애를 진단하기 위한 검사가 아니라 발달지연 가능성을 조기에 확인하고 추가 평가가 필요한 아동을 선별하기 위한 도구로 활용할 필요가 있다.

결론 및 제언

재평가전문위원회는 덴버발달검사가 발달지연 가능성이 있는 아동을 조기에 선별하는 데 활용될 수 있다고 보았다. 다만 발달장애를 확정적으로 진단하는 검사는 아니므로, 검사 결과는 추가 검사 필요성 및 의료진의 판단과 함께 해석할 필요가 있다고 보았다. 또한 실제 의료현장에서의 활용 현황에 대한 추가 근거가 필요하다고 판단하였다.

I 서론

1. 평가배경

덴버발달검사(Denver Developmental Screening Test II, Denver II)는 생후 1개월부터 만 6세까지의 영유아를 대상으로 4개 발달 영역(Gross Motor, Fine Motor-Adaptive, Language, Personal-Social)을 평가하는 선별용 발달검사로, 발달지연 아동을 조기 발견하여 정밀검사나 중재로 연계하는 것을 목적으로 한다.

동 의료기술은 신의료기술평가 제도 도입 이전인 2001년 비급여 항목으로 등재되어 현재 임상에서 사용되고 있다. 이후 동 검사의 사용 현황과 유용성에 대한 근거를 검토할 필요성이 제기되었으며, 2025년 제7차 의료기술재평가위원회(2025. 7. 11.)에서 재평가 계획서 및 소위원회 구성안을 심의한 후 재평가를 수행하였다.

1.1 평가대상 의료기술 개요

1.1.1 덴버발달검사

덴버발달검사(Denver Developmental Screening Test, DDST)는 발달지연 및 장애 위험이 있는 영유아를 선별하기 위해 1967년에 도입되었으며(Frankenburg & Dodds, 1967), 이후 전 세계적으로 널리 활용되었고 15개 이상의 국가에서 표준화되었으며 40여 개 언어로 번역되어 사용되었다(Glascoe & Byrne, 1993). 1992년에는 미국 아동을 대상으로 재표준화하여 Denver II (Denver Developmental Screening Test II)로 개정되었으며, 언어 항목의 확충, 새로운 연령 규준 도입, 행동평가척도의 추가 등 다양한 개선이 이루어졌다(Frankenburg et al., 1992). 개정된 Denver II는 현재까지도 전 세계 임상 및 보건 현장에서 영유아 발달선별검사로 널리 사용되고 있다.

덴버발달검사(Denver II)의 사용대상은 생후 1개월부터 6세 이하의 소아로, 발달지연이 의심되는 아동, 뇌성마비 또는 기타 발달장애를 가진 영유아 등이 해당된다. 본 검사는 유아의 발달 상태를 평가하여 발달 이상 여부를 선별하기 위해 고안된 검사로, 필요시 정밀검사를 통해 발달지연을 조기에 발견하고 개입하는 것을 목적으로 한다(대한소아신경학회, 2022).

Denver II는 개인-사회성, 미세운동-적응, 언어, 전체운동의 4개 발달 영역에 걸친 125개 문항으로 구성되어 있다. 검사자는 아동에게 직접 자극을 제시하거나 보호자 보고를 참고하여 아동의 수행을

평가하며, 각 문항은 P(Pass), F(Fail), R(Refusal), NO(No Opportunity)로 기록한다. 문항 수준에서는 연령선이 75-90% 구간을 지나는 항목에서 실패한 경우 '주의(Caution)', 연령선이 90% 이상 통과한 항목에서 실패한 경우 '지연(Delay)'으로 해석한다. 전체 검사 결과는 지연이 없고 주의가 1개 이하인 경우 '정상(Normal)', 지연이 1개 이상이거나 주의가 2개 이상인 경우 '의심(Suspect)'으로 판정한다. 거부 반응 등으로 판정이 어려운 경우에는 '검사불능(Untestable)'으로 분류하며, 의심 또는 검사불능 결과가 확인된 경우 재검사 또는 정밀평가를 고려한다.

동 검사는 의료인 또는 훈련된 전문가가 시행하며, 아동의 행동과 수행을 직접 관찰해 종합적인 판정을 내린다. 이를 통해 발달지연이 의심되는 아동을 조기에 발견하고, 필요시 정밀검사나 중재로 연계하는 데 활용된다.

표 1.1 발달지연의 정의

*** 발달지연(Developmental delay)** · 아동이 또래에 비해 하나 이상의 발달 단계에 도달하는 속도가 늦은 경우를 의미한다. 발달지연은 다음과 같이 유형별로 구분될 수 있다. - (사고) 능력: 사고, 학습 및 정보 이해가 포함된다. 인지 지연이 있는 아동은 지시를 따르거나 문제를 해결하는 데 어려움을 겪을 수 있다. - 사회성 및 정서적 기술: 다른 사람과 어울리기, 감정 표현하기, 필요를 전달하는 능력 등이 포함된다. 사회적 또는 정서적 지연이 있는 아동은 사회적 신호를 이해하거나 대화를 나누거나 일상의 변화에 대처하는 데 어려움을 겪는다. - 말하기 및 언어 능력: 언어 사용과 이해가 포함된다. 언어 및 언어 지연이 있는 아동은 단어를 말하거나 다른 사람의 말을 이해하는 데 어려움을 겪을 수 있다. - 소근육 및 대근육 운동 능력: 소근육(미세 근육)과 대근육을 조정하는 능력이 포함된다. 소근육 운동 지연이 있는 아동은 손에 물건을 쥐는 데 어려움을 겪거나 색칠하기와 쓰기에 어려움을 겪을 수 있으며, 대근육 운동 지연이 있는 아동은 뒤집기, 앉기, 걷기에 어려움을 겪을 수 있다. · 발달지연은 한 가지 발달 영역에만 영향을 미치는 '고립형 발달지연'이 있을 수 있으며, 발달 영역 중 두 가지 이상에서 현저한 지연이 발생하는 경우에는 '전반적 발달지연'으로 정의된다.

출처: 대한소아신경학회, 2023

1.1.2 영유아건강검진 발달평가 도구로서의 덴버발달검사의 이용

덴버발달검사(Denver II)는 2007년 보건복지부 고시 제2007-125호에서 한국형 유아발달 선별도구(Korean Ages and Stages Questionnaires, K-ASQ)와 함께 영유아건강검진의 발달평가 도구로 채택되었다. 이후 2011년 고시 제2011-166호는 두 도구의 병행 체계를 유지하면서, 연령별 시행 시기와 결과 기록·상담, 의뢰 기준 등의 세부 시행지침을 보완하였다.

그러나 외국의 검사를 표준화한 도구는 우리나라의 정서·사회적인 특성과 문화적 배경을 충분히 반영하지 못하고, 영유아건강검진의 전체 월령을 포괄하지 못하며, 저작권 문제로 웹(web) 상 구현이 어려운 한계가 있었다(보건복지부, 2014). 이에 따라 질병관리본부(현 질병관리청)는 대한소아과학회의 참여하에 2년에 걸쳐 국내 실정에 맞는 한국 영유아 발달선별검사(Korean Developmental Screening Test for Infants and Children, K-DST)를 개발하였으며, 2014년 보건복지부 고시 제2014-111호 개정을 통해 영유아건강검진의 발달평가 도구로 공식 채택되었다. 이로써 덴버발달검사는 국가 영유아건강검진 발달평가 항목에서 제외되었다.

표 1.2 영유아건강검진 발달평가 도구 관련 보건복지부 고시 이력

<table>
<tr><th>관련근거</th><th>내용</th></tr>
<tr><td>보건복지부 고시
제 2007-125호
(2007.12.26.)</td><td>
영유아건강검진실시기준
제2장 건강검진의 시기, 항목 등

제4조(건강검진의 항목 등)
③별표 1의 “발달평가 및 상담”은 공인된 질문 도구를 이용하여 영유아의 발달 사항을 점검하고 상담을 실시하는 것을 말하며, 4개월의 경우 실시하지 않는다.

(별표1)
2. 2차 검진(생후 9~12개월)
<table>
<tr><th>검진항목</th><th>목표질환</th><th>검진비용</th><th>검 사 방 법</th></tr>
<tr><td>2. 발달평가 및 상담</td><td>발달 이상</td><td>6,600원</td><td>○ K-ASQ : 보호자 설문조사 후 의사가 판정하고 상담한다.
- 9~10개월 : K-ASQ 10개월용 사용
- 11~12개월 : K-ASQ 12개월용 사용
○ DENVER II : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.</td></tr>
</table>
3. 3차 검진(생후 18~24개월)
<table>
<tr><th>검진항목</th><th>목표질환</th><th>검진비용</th><th>검 사 방 법</th></tr>
<tr><td>2. 발달평가 및 상담</td><td>발달 이상</td><td>6,600원</td><td>○ K-ASQ : 보호자 설문조사 후 의사가 판정하고 상담한다.
- 9~10개월 : K-ASQ 10개월용 사용
- 11~12개월 : K-ASQ 12개월용 사용
○ Denver II : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.</td></tr>
</table>
</td></tr>
</table>

관련근거	내용

4. 4차 검진(생후 30~36개월)

검진항목	목표질환	검진비용	검 사 방 법
2. 발달평가 및 상담	발달 이상	6,600원	○ 보호자 설문조사 후 의사가 판정하고 상담한다. - 30~31개월 : K-ASQ 30개월용 사용 - 32~34개월 : K-ASQ 33개월용 사용 - 35~36개월 : K-ASQ 36개월용 사용 ○ Denver Ⅱ : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.

5. 5차 검진(생후 54~60개월)

검진항목	목표질환	검진비용	검 사 방 법
2. 발달평가 및 상담	발달 이상	6,600원	○ 보호자 설문조사 후 의사가 판정하고 상담한다. - 54~55개월 : K-ASQ 54개월용 사용 - 56~58개월 : K-ASQ 54, 60개월용 모두 사용하고 의사가 판단 - 59~60개월 : K-ASQ 60개월용 사용 ○ Denver Ⅱ : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.

「건강검진 실시기준」 일부개정
제3장 건강검진의 실시절차 등

제9조(검진 실시시기)
②영유아건강검진은 대상자가 생후 4개월, 9개월, 18개월, 30개월, 42개월, 54개월, 66개월 시기에 각 1회 실시하고, 이 중 영유아 구강검진은 18개월, 42개월, 54개월 시기에 각 1회 실시한다.

제10조(검진 실시절차)
⑤검진기관은 검진 실시에 앞서 다음 각 호의 문진표 등을 구비하여야 하고 대상자는 검진기관의 안내에 따라 해당 문진표 등을 작성하여 제출한다.
5. 영유아건강검진 문진표(별지 제5호서식)
6. 영유아 구강검진 문진표 및 결과통보서(별지 제6호서식)
7. 영유아건강검진 시기별 발달평가 도구(K-ASQ 또는 Denver II)

보건복지부 고시 제2011-166호 (2011. 12. 27.)

(별표4)
5. 5차 검진(생후 42~48개월)

검진항목	목표질환	검진비용	검 사 방 법
2. 발달평가 및 상담	발달 이상	6,600원	○ 보호자 설문조사 후 의사가 판정하고 상담한다. - 42～43개월 : K-ASQ 42개월용 사용 - 44～46개월 : K-ASQ 42개월용, 48개월용 둘 다 사용 후 판단 - 47～48개월 : K-ASQ 48개월용 사용 ○ Denver Ⅱ : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.

관련근거	내용
	6. 6차 검진(생후 54~60개월) <table><tr><th>검진항목</th><th>목표질환</th><th>검진비용</th><th>검 사 방 법</th></tr><tr><td>2. 발달평가 및 상담</td><td>발달 이상</td><td>6,600원</td><td>○ 보호자 설문조사 후 의사가 판정하고 상담한다. - 54~55개월 : K-ASQ 54개월용 사용 - 56~58개월 : K-ASQ 54, 60개월용 모두 사용하고 의사가 판단 - 59~60개월 : K-ASQ 60개월용 사용 ○ Denver II : 의사나 훈련받은 전문가가 실시하고 의사가 상담한다.</td></tr></table> 7. 7차 검진(생후 66~71개월) <table><tr><th>검진항목</th><th>목표질환</th><th>검진비용</th><th>검 사 방 법</th></tr><tr><td>2. 발달평가 및 상담</td><td>발달 이상</td><td>6,600원</td><td>○ 보호자 설문조사 후 의사가 판정하고 상담한다. - 별지 제19호 서식의 영유아 검진용 발달선별 검사도구 및 결과지를 사용한다.</td></tr></table>
보건복지부 고시 2014-111호 (2014. 7. 22.)	**건강검진 실시기준 일부개정** 건강검진 실시기준 일부를 다음과 같이 개정한다. 제10조 제5항 중 제7호 **"영유아건강검진 시기별 발달평가 도구(K-ASQ 또는 Denver II"를 "한국 영유아 발달선별검사 도구(K-DST)"로 한다.**

1.2 국내외 보험 및 행위등재 현황

1.2.1 국내 보험 및 행위등재 현황

덴버발달검사는 현재 비급여로 등재되어 있다.

표 1.3 건강보험 요양급여·비급여 비용 목록 등재 현황(덴버발달검사)

분류번호	코드	분류
		제3부 행위 비급여 목록 제2장 검사료 제3절 기능 검사료 **[신경계기능검사]**
노-693	FZ693	영유아발달검사(한국판 덴버발달검사)
노-694	FZ694	덴버발달검사 Denver Developmental Test

출처: 건강보험심사평가원, 2025

표 1.4 건강보험심사평가원 고시항목 상세

분류번호	노694	EDI코드	FZ694
급여구분	비급여	신의료기술행위	N
행위명(한글)	덴버발달검사		
행위명(영문)	Denver Developmental test		
관련근거	보건복지부고시 제2001-52호(2001.10.26.)		
적용일자	2001-11-01		
적응증	출생 ~ 6세 이내 소아, 발달지연이 의심되는 소아, 뇌성마비, 기타 발달장애질환 영유아 등에서 운동, 사회성, 개인성 등(Gross motor, Fine motor, Language, Personal and social 등)의 발달이 각 소아 연령에 부합되는 지를 판단하여 조기에 치료할 목적으로 시행하는 검사로서 환아에게 여러 가지 자극을 주고 환아의 반응을 관찰하여 기록하는 Screening 검사임		

출처: 건강보험심사평가원 건강보험심사평가원 요양기관업무포털, 의료기준관리, 행위평가신청, 고시항목조회

1.2.2 국외 보험 및 행위등재 현황

미국 행위분류 코드(Current Procedural Terminology, CPT)에서 덴버발달검사만을 위한 CPT 코드는 존재하지 않는다. 다만 선별검사 전반을 위한 코드 중 덴버가 포함된 코드(코드 96110, 96112, 96113)를 확인하였고, 일본의 진료 보수 점수표에는 발달 및 지능 검사(코드 D283)가 등재되어 있는 것으로 확인된다.

표 1.5 국외 보험 및 행위 등재 현황

국가	분류	내용	
미국	CPT*	96110 Developmental screening (eg, developmental milestone survey, speech and language delay screen). with scoring and documentation, per standardized instrument (For an emotional/behavioral assessment, use 96127) (To report developmental testing, see 96112, 96113)	
		96112 Developmental test administration (including assessment of fine and/or gross motor, language, cognitive level, social, memory and/or executive functions by standardized developmental instruments when performed). by physician or other qualified health care professional, with interpretation and report; first hour	
		96113 each additional 30 minutes (List separately in addition to code for primary procedure)	
일본	진료보수 점수표†	D283 Developmental and Intelligence Testing	
		1 Tests that are easy to administer	72 points
		2 Tests that are complex to administer	280 points
		3 Tests that are extremely complex to administer and interpret	450 points

*CPT Professional 2023 and e/M Companion 2023 Bundle
†일본 후생노동성 홈페이지, 2024년판

1.3 국내 이용 현황

동 기술은 비급여 항목으로서 정확한 국내 이용 현황 파악이 어렵다. 국민건강보험 비급여 포털에 따르면, 덴버발달검사의 전국 평균 가격은 58,169원이며, 지역별로 최저 1,600원에서 최고 300,000원까지 큰 차이를 보였다. 덴버발달검사는 주로 '달리 분류되지 않은 증상, 징후와 임상 및 검사의 이상소견' 및 '기타' 질환군에서 활용되었으며, 동일한 질환군에서 사용되었으며, 20세 미만 남성 64.1%, 20세 미만 여성 35.7%의 순으로 이용 빈도가 확인되었다(국민건강보험공단 비급여 정보 포털, 검색일: 2025.10.).

1.4 의학교과서

『소아정신의학(3판)』(홍강의 등, 2023)에 따르면 덴버발달검사는 1967년에 발표되어 전 세계적으로 가장 많이 사용되고 있는 아동 발달선별검사의 하나로, 1992년에 개정 및 재표준화되어 사용되고 있다. 국내에서는 1996년에 한국어판이 표준화되었으며(이근, 1996), 영유아의 발달 상태를 평가하여 발달상 이상 또는 의심 소견이 있는 아동, 즉 발달장애 가능성이 있는 아동을 조기에 선별하는 것을 목적으로 한다고 기술하고 있다.

『재활의학(6판)』(한태륜 등, 2019)과 『소아재활의학(3판)』(대한소아재활·발달의학회, 2021)에서는 덴버발달검사가 출생 시부터 생후 6세 아동을 대상으로 시행할 수 있으며, 간편하고 시행이 쉬운 발달선별검사로 설명하고 있다. 해당 검사는 발달장애를 진단하기 위한 도구가 아니라 발달지연의 가능성이 있는 아동을 선별하는 데 목적이 있다고 기술되어 있다. 검사 항목은 개인-사회성, 소운동-적응, 언어, 대운동의 4가지 영역으로 구성되며, 총 125개의 항목이 각각의 연령군에서 25%, 50%, 75%, 90%의 통과율 기준으로 제시된다. 특정 항목이 동일 연령대의 90% 이상 수행 기준에 미달할 경우 발달지연 가능성을 고려하여 해석하도록 안내하고 있다. 교과서에서는 덴버발달검사가 신뢰도 및 타당도가 보고된 발달평가 도구임을 기술하고 있다.

1.5 관련 국내외 가이드라인

1.5.1 국외 가이드라인

미국소아과학회(American Academy of Pediatrics, AAP)는 2006년 발달감시 및 발달선별에 관한 정책문에서 모든 영유아의 정기 건강관리 방문 시 발달 감시(developmental surveillance)를 시행하고, 생후 9개월, 18개월, 30개월 시점에는 표준화된 발달선별검사(standardized developmental screening test)를 시행하도록 권고하였다. 해당 문헌에서는 발달선별검사가 진단이나 치료계획을 도출하는 검사가 아니라, 아동의 발달이 또래 기준과 차이를 보이는 영역을 확인하는 절차라고 설명하고 있다. 또한 선별검사 결과가 우려되는 경우에는 발달평가 및 의학적 평가로 연계해야 하며, 발달 감시 또는 선별을 통해 발달장애 고위험으로 확인된 경우 진단적 발달평가가 필요하다고 제시하였다(AAP, 2006).

AAP 2006년 정책문은 특정 선별도구를 단일하게 권고하지는 않았으나, 발달선별도구의 예시 목록에 연령 및 단계 질문지(Ages and Stages Questionnaires, ASQ), 부모 발달상태 평가(Parents' Evaluation of Developmental Status, PEDS), Denver-II 등을 포함하였다. 이 중 Denver-II는 표현 및 수용 언어(expressive and receptive language), 대근육운동(gross motor), 미세운동(fine motor), 개인-사회성(personal-social skills) 영역을 평가하는 직접 시행 방식의 발달선별도구로 설명되어 있다.

1.5.2 국내 가이드라인

국내에서는 덴버발달검사(Denver II)의 활용과 관련한 공식적인 임상진료지침 또는 학회 차원의 가이드라인은 현재까지 확인되지 않는다.

1.6 관련 체계적 문헌고찰

"Denver Developmental Screening Test" OR "Denver II"를 검색어로 하여 PubMed에서 검색한 결과, 체계적 문헌고찰 1편(Rah et al., 2023)이 확인되었다.

Rah 등(2023)은 아동 발달선별검사의 실제 임상현장에서의 진단 정확도를 체계적으로 검토하여 메타분석을 수행한 연구로, 총 56편의 연구를 분석하였다. 이 연구에서 가장 빈번하게 사용된 선별검사는 ASQ, DDST, PEDS 였으며, 덴버발달검사를 이용한 연구는 8편이었다. 전체적으로 선별검사의 통합 민감도는 0.75 (95% CI 0.69-0.80), 특이도는 0.76 (95% CI 0.71-0.80)으로 보고되었다.

표 1.6 관련 체계적 문헌고찰

<table>
<tr><th>제목</th><th>Systematic Review and Meta-analysis: Real-World Accuracy of Children's Developmental Screening Tests</th></tr>
<tr><td>제1저자(연도)</td><td>Rah (2023)</td></tr>
<tr><td>국가</td><td>한국</td></tr>
<tr><td>연구목적</td><td>전 세계 임상 환경에서 사용되는 발달선별검사 목록을 제시하고 그 정확성에 대한 광범위한 추정치 제공</td></tr>
<tr><td>데이터베이스</td><td>• PubMed, PsycINFO, Cochrane, EMBASE, CINAHL, Google Scholar databases up to December 2020
• Medical Subject Heading terms and related free-text synonyms for the DD subconditions in addition to "developmental disabilities"</td></tr>
<tr><td>PICO</td><td>• 대상환자: 13세 미만의 아동
• 중재(intervention): 발달선별검사(ASQ, DDST, PEDS 등)
• 비교중재(comparators): 발달 진단검사 (Bayley Scales of Infant Development 등)
• 결과변수: 발달선별검사의 진단 정확도 (민감도, 특이도 등)</td></tr>
<tr><td>연구결과</td><td>- 선택 연구: 56개 연구
- 덴버발달검사 포함: 8편
<table>
<tr><th>Study</th><th>N</th><th>Age (mon)</th><th>대상아동</th><th>선별 검사</th><th>민감도</th><th>특이도</th><th>DOR</th></tr>
</table></td></tr>
</table>

제목	Systematic Review and Meta-analysis: Real-World Accuracy of Children's Developmental Screening Tests							
	De Kleine (2006)	396	60	미숙아/ 저체중아	DDST	0.52 [0.45-0.60]	0.80 [0.75-0.85]	4.49 [2.88-7.01]
	Glascoe & Byrne (1993)	89	7-70	보육시설 재원 아동	DDST, BDIST	0.83 [0.59-0.96]	0.46 [0.35-0.59]	4.34 [1.15-16.33]
	Glascoe (1992)	104	3-72	보육시설 재원 아동	DDST	0.83 [0.59-0.96]	0.43 [0.32-0.54]	3.78 [1.02-14.01]
	Grant and Gittelsohn (1972)	298	42-54	지역사회 일반 아동	DDST	0.79 [0.61-0.91]	0.58 [0.52-0.64]	5.07 [2.13-12.10]
	Lichtenstein (1981)	40	43-59	일반 아동	DDST	0.33 [0.01-0.91]	0.92 [0.778-0.98]	5.67 [0.39-82.24]
	Mirrett (2004)	18	9-18	Fragile X Syndrome 아동	DDST, BDIST	0.86 [0.57-0.98]	1.00 [0.16-1.00]	1.93 [0.06-62.17]
	Palti and Tepper (1987)	176	6-24	일반 아동	DDST	0.07 [0.01-0.24]	0.99 [0.95-1.00]	5.62 [0.76-41.65]
	Sondaar (2007)	368	N/I	미숙아	DDST	0.72 [0.59-0.83]	0.79 [0.74-0.83]	9.64 [5.17-17.97]
결론	• 가장 자주 사용된 선별 검사는 ASQ, DDST, PEDS 였음 • 발달선별검사의 진단 정확도는 민감도와 특이도 평균 70% 이상으로 나타났으나, 실제 임상에서의 효과성은 발달 우려 여부 등 다양한 요인에 따라 달라질 수 있음 • 따라서 선별검사의 유효성을 확보하기 위해서는 표준화된 진단 정확도 검증 연구 설계가 필요함							

ASQ, Ages and Stages Questionnaire ; DDST, Denver Developmental Screening Test ; PEDS, Parents' Evaluation of Developmental Status

1.7 관련 의료기술

덴버발달검사(Denver II)는 영유아의 전반적인 발달 상태를 관찰하여 발달지연 가능성이 있는 아동을 조기에 선별하기 위한 검사도구로, 발달장애의 진단이나 확정 평가를 목적으로 개발된 도구는 아니다. 따라서 선별검사 결과에서 발달지연이 의심되거나 우려되는 소견이 확인된 경우에는 검사 결과를 단독으로 해석하기보다, 평가 목적과 임상 상황에 따라 정밀 발달평가 또는 영역별 전문 평가로 연계하는 것이 필요하다(AAP, 2006).

영유아 발달평가에서는 평가 목적과 임상 상황에 따라 선별검사, 정밀 발달평가, 영역별 전문 평가가 순차적 또는 병행적으로 활용될 수 있다. 이 중 베일리영아발달측정(Bayley Scales of Infant and Toddler Development, Bayley)반면 K-DST는 국가 영유아건강검진에서 발달지연 위험군을 선별하기 위해 활용되는 보호자 보고 기반의 발달선별검사이다(Chung et al., 2020). 따라서 덴버발달검사, Bayley 검사, K-DST는 모두 영유아의 발달 상태를 평가한다는 점에서 연관성이 있으나, 검사 목적, 시행 방식, 적용 환경에는 차이가 있다. 덴버발달검사는 임상 현

장에서 관찰 기반의 1차 선별검사로 활용되고, Bayley 검사는 선별 이후 발달 수준을 보다 구체적으로 평가하기 위한 정밀 발달평가 도구로 활용되며, K-DST는 국가건강검진 체계에서 보호자 보고 기반의 표준 발달선별검사로 활용된다. 이들 도구는 상호 대체 관계라기보다, 발달선별–정밀 평가–추적 및 중재 연계라는 평가 흐름 속에서 서로 보완적인 위치에 있다. 이들 검사도구의 기능적 위치와 상호 관계를 설명하기 위해, 영유아 발달평가 검사 활용의 전반적 흐름을 도식화하여 제시하였다(표 1.7).

표 1.7 영유아 발달평가 검사의 활용 예시

전반적 발달선별(Screening)
K-DST / Denver II
정밀발달평가 또는 영역별 선택 검사
• Bayley-III/IV(인지·언어·운동·사회정서·적응행동) • 영역별 전문 진단 도구 · 언어: PRES, REVT, SELSI 등 · 운동·조정: PDMS-2, AIMS 등 · 인지·지능: WPPSI-IV, WISC-V 등 · ASD 평가: ADOS-2, ADI-R, CARS-2 · 정서·행동: CBCL, BASC-3 등 · 자조·적응행동: Vineland-3 등
경과 추적·기능 평가
발달재평가, 치료경과 확인, 중재 연계 결정

· Screening
K-DST: Korean Developmental Screening Test for Infants and Children
Denver II : Denver Developmental Screening Test II
· 다영역 정밀 진단
Bayley-III/IV: Bayley Scales of Infant and Toddler Development (3rd/4th Edition)
· 언어 진단
PRES: Preschool Receptive-Expressive Language Scale
REVT: Receptive & Expressive Vocabulary Test
SELSI: Sequenced Language Scale for Infants
· 운동·조정 진단
PDMS-2: Peabody Developmental Motor Scales, Second Edition
AIMS: Alberta Infant Motor Scale
· 인지·지능 진단
WPPSI-IV: Wechsler Preschool and Primary Scale of Intelligence – IV
WISC-V: Wechsler Intelligence Scale for Children – V
K-ABC-II: Kaufman Assessment Battery for Children – II
ASD (Autism Spectrum Disorder) 진단
ADOS-2: Autism Diagnostic Observation Schedule – 2
ADI-R: Autism Diagnostic Interview – Revised
CARS-2: Childhood Autism Rating Scale -2
· 정서/행동 진단
K-CBCL / CBCL: Korean Child Behavior Checklist / Child Behavior Checklist
BASC-3: Behavior Assessment System for Children – Third Edition
· 자조·적응행동
Vineland-3: Vineland Adaptive Behavior Scales, Third Edition

〈표 1.7〉은 영유아 발달평가가 일반적으로 전반적 발달선별(screening) 단계에서 시작하여, 필요시 정밀 발달평가 또는 영역별 전문 평가, 그리고 결과 추적 및 중재 연계로 이어지는 과정을

개념적으로 나타낸 것이다. 이 흐름도는 특정 검사도구를 필수적으로 시행해야 하는 단계별 기준을 제시하기 위한 것이 아니라, 실제 임상 및 보건 현장에서 관찰되는 발달평가의 일반적 활용 맥락을 이해하기 위한 예시적 구조이다. 실제 진료 현장에서는 아동의 연령, 기능 수준, 협조 가능성, 의심되는 문제 영역, 기관의 자원 및 인력 여건 등에 따라 평가 도구의 선택과 조합이 달라질 수 있다.

따라서 〈표 1.7〉은 영유아 발달평가 과정에서 덴버발달검사를 포함한 주요 검사도구들의 기능적 위치와 상호 관계를 정리하기 위한 개념적 흐름도로, 이후 제시되는 Bayley 검사와 K-DST의 역할 및 적용 범위를 설명하기 위한 기준 틀로 제시하였다.

1.7.1 베일리영아발달측정

베일리영아발달측정(Bayley Scales of Infant and Toddler Development, Bayley)은 생후 1개월부터 42개월(약 3세 6개월)까지의 영유아를 대상으로 발달 수준을 정밀하게 평가하기 위해 개발된 표준화된 발달평가 도구이다. 본 검사는 발달지연이나 장애가 의심되는 아동을 대상으로, 인지·언어·운동· 사회정서·적응행동 등 발달 전반을 정량적으로 평가하는 데 활용된다(Bayley, 2006).

(1) 발달 영역 구성 및 판별 체계

Bayley-II에서는 정신발달지수(Mental Development Index, MDI), 정신운동발달지수(Psychomotor Development Index, PDI), 그리고 행동평가척도(Behavior Rating Scale, BRS)의 세 가지 영역으로 구성되어 있으며, 각각 인지·언어 발달, 대·소근육 운동, 검사 수행 시 행동 특성을 평가한다(Bayley, 1993). 이후 개정된 Bayley-III에서는 평가 체계가 확장되어, 인지(Cognitive), 언어(Language; 수용 및 표현), 운동(Motor; 대·소근육), 사회정서(Social- Emotional), 적응행동(Adaptive Behavior)의 다섯 영역을 포함한다. 각 영역별 점수는 소척도점수(Scaled Score) 및 종합표준점수(Composite Standard Score)로 산출되며, 발달지연 정도를 객관적 수치로 제시한다(Bayley, 2006).

(2) 검사 방법 및 소요 시간

검사는 훈련된 전문가가 아동에게 발달 수준에 맞는 놀이 기반 과제를 제시하고, 영유아의 반응과 수행을 관찰·채점하는 방식으로 시행된다. Bayley-III의 직접시행 영역은 인지 91문항, 언어 97문항(수용언어 49문항, 표현언어 48문항), 운동 138문항(소근육 66문항, 대근육 72문항)으로 구성되어 총 326문항이다. 사회정서 및 적응행동 영역은 보호자 설문을 통해 별도로 평가된다. 실제 검사는 아동의 연령과 수행 수준에 따라 시작점과 중지 기준을 적용하므로 모든 문항을 일괄적으로 시행하지는 않는다(Bayley, 2006). 검사 소요 시간은 보통 45-60분 정도이나, 아동의 연령, 협조도, 발달 수준 및 시행 범위에 따라 최대 90분 내외까지 소요될 수 있다(Bayley, 2006; Ranjitkar et al., 2018).

(3) 임상적 활용 및 급여 현황

Bayley 검사는 조산아, 저체중아 등 고위험 영유아의 발달평가 및 추적관리, 발달지연 또는

발달장애가 의심되는 아동의 정밀 발달평가, 중재 전후 발달 수준 평가 등에 활용된다(Johnson & Marlow, 2006; Jeong et al., 2017). 현재 건강보험 급여 항목(행위코드: FZ695, 분류번호: 나-627)으로 등재되어 있으며, 상대가치점수는 2,084.2점, 의원급 196,120원, 병원급 171,320원으로 책정되어 있다.

표 1.8 건강보험 요양급여·비급여 비용 목록 등재 현황(베일리영아발달측정)

분류번호	코드	분류	점수
		제1편 행위 급여·비급여 목록 및 급여 상대가치점수 제2부 행위 급여 목록·상대가치점수 및 산정지침 제2장 검사료 제3절 기능 검사료 **[신경계기능검사]**	
나-627	FZ695	베일리영아발달측정 Bayley Scales of Infant Development	2,084.20

출처: 건강보험심사평가원. 2025

표 1.9 건강보험심사평가원 행위기술서(베일리영아발달측정)

분류번호	나627	EDI코드	FZ695
급여구분	급여	신의료기술행위	N
행위명(한글)	베일리영아발달측정		
행위명(영문)	Bayley Scales of Infant Development		
적응증	[적응증] 1. 영유아 발달에 문제가 의심되는 경우 2. 발달선별검사에서 이상이 있는 경우 [실시방법] ○ 시술전 1. 문진한다. (아동 전반적인 발달에 대해서 문진) 1) 의학적인 신경학적인 검사를 포함하여 대운동, 소운동, 인지, 적응 행동, 사회성 발달 정도를 포함하여 양육 환경, 교육 및 치료 환경에 대해서 정보를 습득한다. 2. 부모에게 베일리 검사에서 요구하는 행동 관찰에 대한 설문지를 배부한다. ○ 시술 중 1. 발달재활과정을 숙지하고 상기 검사에 훈련을 받은 보건의료인이 아동과 1:1로 표준화된 검사 도구 세트를 사용하여 검사 시행한다. 2. 검사지에 세부검사항목 통과여부를 기록한다. 3. 발달 상태를 인지척도, 동작척도, 행동평정척도를 이용하여 평가한다. 4. 부모의 설문지를 받아 기록한다. 5. 점수를 계산하고 검사 시의 특이사항이나 통과, 비통과 항목에 대한 의견과 각기 세부내용을 빠짐없이 리포트로 작성한다. ○ 시술 후 1. 사전에 검사결과지 판독 및 해석하여 임상양상과 합당하게 평가가 되었는지 확인한다. 2. 검사 결과를 설명한다. 3. 현재 아동이 발달장애가 있는지, 향후 예후 및 검사 계획 설명, 현재 필요한 치료계획에 대해서 설명 및 상담한다.		

출처: 건강보험심사평가원 요양기관업무포털

표 1.10 베일리영아발달측정 관련 보건복지부 고시 이력

<table>
<tr><th>관련근거</th><th>내용</th></tr>
<tr><td rowspan="2">보건복지부 고시
제2016-212호
(2016.11.18.)</td><td>「건강보험 행위 급여·비급여 목록표 및 급여 상대가치점수」 일부개정

제1편 제2부 제2장 제3절 [신경계기능검사]에 '나-627'란을 다음과 같이 신설한다.
<table>
<tr><th>분류번호</th><th>코드</th><th>분 류</th><th>점 수</th></tr>
<tr><td>나-627</td><td>FZ695</td><td>[신경계기능검사]
베일리영아발달측정
Bayley Scale of Infant Development</td><td>1,812.35</td></tr>
</table></td></tr>
<tr><td>보건복지부 고시
제 2016-275호
(2016.12.29.)</td><td>「요양급여의 적용기준 및 방법에 관한 세부사항」일부개정

Ⅰ. 행위 일반사항, 제1장, 제2장, 제15장, 제19장, Ⅶ. 완화의료 제2부 중 항목, 제목 및 세부인정사항 일부를 각각 별지 1과 같이 신설한다.

제2장 검사료
<table>
<tr><th>항목</th><th>제목</th><th>세부인정사항</th></tr>
<tr><td>나627 베일리영아발달측정</td><td>베일리영아발달측정 수가산정방법</td><td>베일리영아발달측정은 영유아 발달검사나 영유아건강검진상 발달지연이 의심되어 시행하는 경우에 다음과 같이 급여함.

- 다 음 -

가. 생후 24개월까지 : 최대 3회
나. 생후 24개월 초과~42개월 : 연 1회
다. 생후 42개월을 초과하더라도 발달 수준이 42개월 미만으로 확인되는 경우에는 사례별로 인정함</td></tr>
</table></td></tr>
</table>

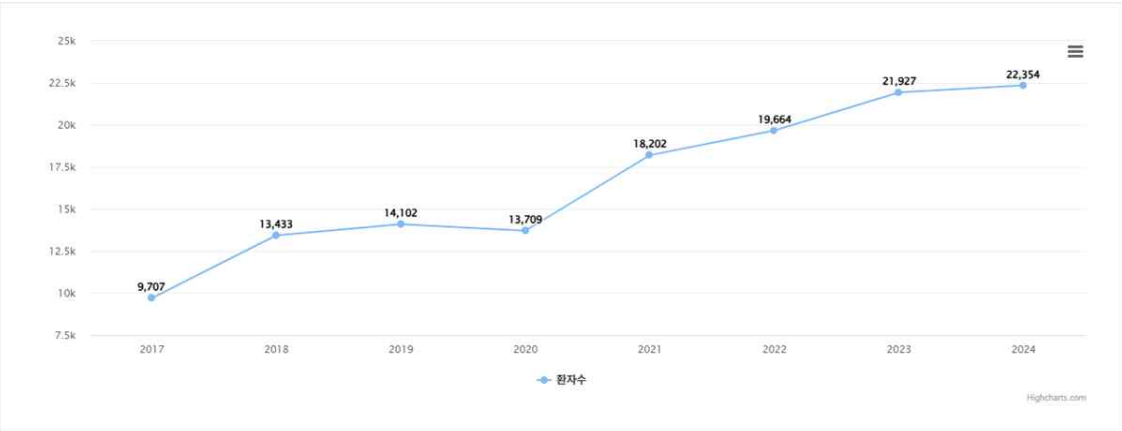

그림 1.1 베일리영아발달측정 환자 수 추이

(출처: HIRA 빅데이터개방포털 홈페이지)

표 1.11 상대가치점수

수가코드	분류번호	명칭	급여구분	단가	
				의원	병원
FZ695	나627	베일리영아발달측정	급여	196,120	171,320

출처: 건강보험심사평가원 홈페이지 (검색일: 2025.10.26.)

표 1.12 베일리영아발달측정 이용 현황

코드	구분	2020년	2021년	2022년	2023년	2024년
FZ695	환자 수(명)	13,709	18,202	19,664	21,927	22,354
	총사용량, 회	14,993	19,902	21,542	24,150	24,823
	진료 금액, 천원	1,642,087	3,565,808	3,913,512	4,454,867	4,701,393

출처: HIRA빅데이터개방포털 홈페이지(검색일: 2025.10.26.)

1.7.2 한국 영유아 발달선별검사(K-DST)

(1) 개요

한국 영유아 발달선별검사(Korean Developmental Screening Test for Infants and Children, K-DST)는 우리나라 영유아의 발달 특성과 사회문화적 배경을 반영하여 개발된 국가 표준 발달선별검사 도구이다. 이는 기존 외국 도구 기반 검사(Denver II, K-ASQ)가 한국 아동의 언어·사회문화적 특성을 충분히 반영하지 못하고, 저작권 및 전산화 활용 제약 등의 한계를 보완하기 위해 고안되었다(보건복지부, 2014a).

K-DST는 2012-2014년에 국내 4-71개월 아동을 대상으로 표준화 및 타당화 과정을 거쳐 개발되었고, 2014년 보건복지부 고시 제2014-111호에 의해 영유아건강검진의 발달평가 도구로 공식 채택되었다(보건복지부, 2014a; 보건복지부, 2014b). 현재 K-DST는 국가 영유아건강검진의 발달평가 항목으로 사용되고 있으며, 영유아의 발달지연 위험군을 조기에 선별하고 필요한 경우 추가 평가로 연계하기 위한 목적으로 활용된다.

(2) 검사 구성 및 실시 방법

K-DST는 생후 4개월부터 71개월(약 6세)까지의 영유아를 대상으로, 발달단계별로 구분된 총 5~6개의 발달영역을 평가한다. 대근육운동(Gross motor), 소근육운동(Fine motor), 인지(Cognition), 언어(Language), 사회성(Sociality), 자조(Self-help)의 여섯 개 발달 영역으로 구성되며, 각 영역은 8문항씩 포함한다. 자조 영역은 18개월 이상 아동에게 적용되므로, 연령 구간에 따라 총 40~48문항으로 구성된다(질병관리본부, 2017; Chung et al., 2020).

K-DST는 부모 또는 주요 양육자가 직접 응답하는 부모작성형 검사로, 영유아의 일상적 행동 관찰에 근거한 평가를 가능하게 한다. 각 문항은 '전혀 할 수 없다', '하지 못하는 편이다', '할 수 있는 편이다', '잘 할 수 있다'의 4점 척도로 응답하며, 각 영역별 점수를 합산하여 월령별 기준점수와 비교한다. 검사 결과는 빠른 수준, 또래 수준, 추적검사 요망, 심화평가 권고 등으로 분류되며, 문항 응답에는 대체로 5~10분 정도가 소요된다(질병관리본부, 2017). 필요시 심화평가 권고군은 전문의 진료

또는 정밀 발달평가로 연계된다.

(3) 신뢰도 및 타당도

Chung et al.(2020)에 따르면, K-DST 초판은 2012~2014년에 4~71개월 아동 3,284명을 대상으로 표준화 및 타당화 과정을 거쳐 개발되었고, 이후 2015~2016년 영유아건강검진 참여자 자료를 기반으로 재표준화 및 재타당화가 수행되었다. 개정판에서 내적일관성은 0.73~0.93, 검사-재검사 신뢰도는 0.77~0.88로 보고되었으며, 발달지연 선별에 대한 민감도는 0.833, 특이도는 0.979로 제시되었다(Chung et al., 2020).

(4) 활용 및 제도적 현황

K-DST는 현재 국가 영유아건강검진의 표준 발달선별검사로 사용되고 있으며, 보건복지부 고시 제2014-111호에 따라 기존 Denver II와 K-ASQ를 대체하여 도입되었다(보건복지부, 2014a; 보건복지부, 2014b). K-DST는 생후 4~71개월 영유아를 대상으로 하는 발달평가에 활용되며, 영유아건강검진 월령주기인 생후 4~6개월, 9~12개월, 18~24개월, 30~36개월, 42~48개월, 54~60개월, 66~71개월 시기에 아동의 발달 상태를 확인하는 데 사용된다. 검사비용은 영유아건강검진 제도 내에서 국민건강보험공단이 부담한다.

K-DST는 선별 목적의 도구로, 발달장애를 확정 진단하기 위한 검사가 아니라 발달지연 위험군을 조기에 발견하고 필요한 경우 심화평가, 전문의 진료 및 조기 개입으로 연계하는 것을 목표로 한다(질병관리본부, 2017; Chung et al., 2020).

표 1.13 영유아 발달검사의 목적 및 검사 특성 비교

구분	K-DST	Denver II / 한국판 덴버발달검사	Bayley III
목적	국가 영유아건강검진용 발달선별	임상 선별 및 필요시 추적관찰	정밀 발달평가 및 진단
대상	4~71개월 영유아	0~6세 영유아 및 아동	생우 1~42개월 영유아, 주로 발달지연 의심 또는 고위험군
시행자	보호자 설문 응답 + 의료인 판정	훈련된 검사자 (의료인, 치료사 등)	훈련된 전문가 (임상심리사, 치료사 등)
검사형태	보호자 보고식 설문	아동 과제 수행 + 검사자 관찰·기록	표준화된 과제 직접 수행 및 보호자 설문
결과/점수	영역별 선별 결과 및 권고	통과, 주의, 지연에 따른 정상/의심/검사불능 등 판정	소척도점수 및 종합 표준점수
비용	국가검진 내 본인부담 없음	비급여	급여
검사시간	약 10-10분(보호자 응답 기준), 검진 판정·상담 포함 시 10분 내외	약 15-20분	보통 약 45~60분, 연령·협조도·시행범위에 따라 최대 90분 내외
활용	영유아건강검진 (보건소, 일차의료기관)	소아청소년과, 재활의학과, 발달클리닉, 보건소 등	병원, 대학병원, 전문클리닉, 연구기관 등
장점	전국 단위 선별에 적합하고 접근성 이 높음	직접 관찰 기반으로 영역별 발달 수행을 확인할 수 있음	표준점수를 통해 발달 수준을 정량화하고 세부 영역별 평가가 가능함
한계	보호자 보고 편향 가능성이 있으며 확진 도구가 아님	표준점수 산출이 제한적이며 검사자 훈련이 필요함	시간·비용 부담이 크고 전문 인력이 필요함

출처: 질병관리본부(2017), Chung et al.(2020), Frankenburg et al.(1992), Shin et al.(2005), Bayley(2006), Ranjitkar et al.(2018).

베일리 영아발달검사는 생후 42개월(3세 6개월)까지의 영유아를 대상으로 하며, 그 이후 연령에서는 사용이 제한된다. 42개월 이상의 아동에게 Bayley 검사를 적용할 경우 검사 문항이 연령 수준에 비해 지나치게 쉬워, 점수가 상한치에 몰리는 '천장효과(ceiling effect)'가 발생할 수 있다. 이로 인해 실제 발달 수준의 세부적 차이를 구분하기 어렵게 되며, 발달지연 여부의 변별력이 감소한다.

따라서 42개월 이상의 영유아에서는 대근육 운동을 포함한 전반적 발달을 평가할 수 있는 대표적 선별검사로 덴버발달검사가 활용되며, 필요시 기타 보조적 평가 도구를 병행하여 종합적으로 판단한다.

표 1.14 발달검사의 연령별 적용 가능 구간

연령(개월)	K-DST	Denver II	Bayley III	비고
3개월	X	가능	가능	K-DST는 국가 영유아건강검진의 발달선별검사 적용 월령 전 단계이다. Denver II는 0~6세 아동을 대상으로 하며, Bayley-III는 생후 16일~42개월까지 적용 가능하다.
4-42개월	가능	가능	가능	K-DST는 4~71개월 영유아를 대상으로 하며, 국가 영유아건강검진 월령 구간에 따라 시행된다. Denver II와 Bayley-III도 해당 연령에서 적용 가능하다.
43-71개월	가능	가능	X	K-DST는 71개월까지 적용 가능하고 Denver II도 0~6세 범위에 포함되나, Bayley-III는 표준 적용 연령 범위를 벗어난다.
72개월(6세)	X	가능	X	Denver II는 0~6세 아동을 대상으로 하므로 적용 가능하나, K-DST의 국가검진 적용 월령과 Bayley-III의 적용 연령 범위를 벗어난다.

· 주: K-DST는 국가 영유아건강검진 월령 구간인 4~6, 9~12, 18~24, 30~36, 42~48, 54~60, 66~71개월에 시행됨
· 출처: 보건복지부(2014b), 질병관리본부(2017), Chung et al.(2020), Frankenburg et al.(1992), Bayley(2006).

2. 평가목적

본 평가의 목적은 발달지연이 의심되는 영유아에서 덴버발달검사의 임상적 유용성 등을 종합적으로 검토하여 향후 관련 의사결정에 활용할 수 있는 근거를 제공하고자 하였다.

II 평가방법

1. 문헌고찰

1.1 개요

본 평가에서는 덴버발달검사의 임상적 유용성을 평가하기 위하여 문헌고찰을 수행하였다. 동 검사는 국내외에서 장기간 활용되어 온 발달선별검사로, 본 평가에서는 검사도구의 진단정확도를 재평가하기보다는 검사 특성, 임상적 활용 현황 및 실제 사용 양상을 종합적으로 검토하는 데 중점을 두었다. 이에 따라 데이터베이스 검색과 수기검색을 병행하여 관련 문헌을 포괄적으로 검토하였다. 구체적인 평가방법은 평가목적을 고려하여 "덴버발달검사 소위원회(이하 '소위원회'라 한다)"의 검토에 따라 최종 확정하였다.

1.2 핵심질문

본 평가의 핵심질문은 "덴버발달검사는 발달지연이 의심되는 소아를 선별하는데 임상적으로 유용성이 있는가?"이다.

1.3 연구검색

1.3.1 국외

국외 연구검색은 Ovid MEDLINE, APA PsycInfo(via Ovid) 2개의 전자데이터베이스를 이용하였다. 검색어는 Ovid MEDLINE에서 사용된 검색어를 기본으로 각 자료원의 특성에 맞게 수정하였으며 MeSH term, 논리연산자, 절단검색 등의 검색기능을 적절히 활용하였다.

표 2.1 국외 전자데이터베이스

국외 연구 검색원	URL 주소
Ovid MEDLINE	http://ovidsp.ovid.com
APA PsycInfo(via Ovid)	http://ovidsp.ovid.com

1.3.2 국내

국내 연구검색은 코리아메드(KoreaMed)를 이용하여 수행하였다. 본 연구에서는 국외 검색에 활용한 검색전략을 기본 틀로 삼되, 각 데이터베이스의 지원 기능에 따라 이를 조정하여 적용하였다. 논리연산자나 절단검색과 같이 일부 데이터베이스에서 제공되지 않는 기능은 적절히 수정하거나 간소화하였으며, 데이터베이스별 특성을 고려하여 영문과 국문 검색어를 혼용하여 사용하였다.

표 2.2 국내 전자데이터베이스

국내 연구 검색원	URL 주소
KoreaMed	http://koreamed.org

1.3.3 수기 검색

전자검색원의 검색 한계를 보완하기 위하여 선행 체계적 문헌고찰 및 연구 검색과정에서 확인되거나 본 연구 주제와 관련된 참고문헌 등을 토대로, 본 연구의 선택/배제 기준에 적합한 연구를 추가로 검토하여 선정 여부를 판단하였다.

1.4 연구선정

1차 연구선정 단계에서는 제목 및 초록을 바탕으로 본 평가의 평가대상과 관련성이 없는 것으로 판단되는 연구를 배제하고, 2차 단계에서는 연구의 전문(full-text)을 검토하여 본 평가의 선택기준에 맞는 연구를 최종적으로 선정하였다. 구체적인 연구의 선택 및 배제 기준은 〈표 2.3〉과 같다.

표 2.3 연구의 선택 및 배제기준

선택기준(Inclusion Criteria)	배제기준(Exclusion Criteria)
- 6세 이하의 발달지연이 의심되는 소아를 대상으로 수행된 연구 - 정상 아동, 발달지연 의심군, 고위험군(조산아/저체중아), 특정 질환군 등 다양한 집단 연구 포함 - 덴버발달검사가 수행된 연구	- 덴버발달검사가 아닌 다른 도구만을 평가한 연구 - 사전에 정의한 의료결과가 하나 이상 보고되지 않은 경우 - 인간 대상 연구가 아닌 경우(동물연구 또는 전임상연구) - 원저가 아닌 연구(종설, letter, comment 등) - 회색문헌(초록만 발표된 연구, 학위논문, 기관보고서 등 peer-review를 거치지 않은 경우) - 중복 출판된 문헌: 대상자가 중복되고, 보고된 결과지표도 동일한 연구

1.5 자료정리 및 분석

최종 선택된 문헌을 대상으로 연구대상, 검사도구, 비교검사 또는 참고표준, 주요 결과지표 및 연구결과를 추출하여 정리하였다. 본 평가에서는 덴버발달검사의 임상적 유용성을 확인하기 위하여 선택연구의 결과를 검사 도구의 타당도 및 선별정확도와 임상적 활용 양상으로 구분하여 검토하였다.

검사 도구의 타당도 및 선별정확도는 참고표준 또는 비교검사와의 관련성, 민감도, 특이도, 양성예측도, 전체 분류 정확도, 발달지수와의 상관성 등을 중심으로 정리하였다. 임상적 활용 양상은 검사 대상자의 특성, 검사 수행 환경, 발달지연 의심 아동의 선별, 정밀평가 의뢰 또는 중재 연계 여부 등을 중심으로 검토하였다.

선택연구의 연구설계, 대상 집단, 결과지표 및 보고 방식에 이질성이 있어 정량적 합성은 수행하지 않았으며, 문헌에서 보고된 결과를 중심으로 질적 검토를 수행하였다. 또한 문헌고찰 결과만으로 확인하기 어려운 국내 임상현장에서의 실제 활용 현황과 보건의료체계 내 적용 맥락을 보완적으로 확인하기 위하여 전문가 자문을 추가로 실시하고, 그 의견을 함께 검토하였다.

2. 전문가 의견

2.1 개요

덴버발달검사는 선별검사라는 도구 특성과 비급여 기반의 운영 구조, 진료과 및 의료기관별 적용 방식의 차이로 인해 국내 임상현장에서의 실제 사용 현황과 활용 양상을 문헌만으로 충분히 확인하는 데 한계가 있었다. 이에 본 평가에서는 문헌고찰 결과를 보완하고 국내 임상현장에서의 활용 현황 및 보건의료체계 내 적용 맥락을 확인하기 위하여 관련 분야 전문가를 대상으로 추가 자문을 실시하였다.

2.2 전문가 추가 자문

추가 자문에서는 덴버발달검사의 실제 사용 현황, 적용 목적, 검사 결과의 활용 방식 및 국내 보건의료 환경에서의 역할 등에 대한 의견을 수렴하였다.

2.3 검토 내용

전문가 추가 자문에서는 덴버발달검사의 국내 임상현장 내 활용 양상과 검사 결과의 해석상 고려사항을 중심으로 검토하였다.

주요 검토 내용은 다음과 같다.

- 동 검사가 발달지연이 의심되는 영유아를 조기에 선별하기 위한 도구로 실제 임상현장에서 어떻게 활용되고 있는지
- 검사 결과가 정밀 발달평가 의뢰, 추적관찰, 발달재활서비스 또는 중재 연계 과정에서 어떠한 역할을 하는지
- 한국 영유아 발달선별검사(K-DST) 등 기존 발달선별검사와의 관계에서 덴버발달검사의 활용 의미
- 검사 결과의 해석 시 고려사항 및 선별검사로서의 한계
- 진료과 및 기관별 활용 양상과 실제 사용 현황 파악의 제한점

2.4 의견 수렴 및 반영 방법

관련 전문가의 추가 자문 의견은 문헌고찰 결과에서 확인하기 어려운 국내 임상현장의 활용 현황과 보건의료체계 내 적용 맥락을 보완적으로 확인하기 위한 자료로 활용하였다.

소위원회에서는 체계적 문헌고찰 결과와 전문가 추가 자문 의견을 함께 검토하여 덴버발달검사의 임상적 활용 현황, 검사 결과의 해석 및 국내 적용 상황에 대해 논의하였으며, 이를 평가결과 및 결론 도출 과정에 참고하였다.

3. 결론 도출

본 평가는 덴버발달검사의 임상적 유용성을 확인하기 위하여 체계적 문헌고찰 결과를 중심으로 검토하였으며, 관련 전문가의 추가 자문 의견을 참고하여 국내 임상현장에서의 활용 현황과 적용 맥락을 함께 고려하였다.

소위원회에서는 문헌고찰 결과와 전문가 추가 자문 의견을 종합적으로 검토하여 결론을 도출하였다. 결론 도출 시에는 동 검사가 발달지연 위험군을 조기에 선별하기 위한 초기 평가 도구로 활용될 수 있는지, 검사 결과가 정밀평가 의뢰 또는 추적관찰 대상자를 구분하는 데 참고자료로 사용될 수 있는지, 국내 임상현장에서의 활용 범위와 제한점이 무엇인지를 종합적으로 검토하였다.

또한 동 검사는 장애를 확정적으로 진단하기 위한 도구가 아니라 발달지연 가능성이 있는 영유아를 선별하기 위한 검사이므로, 검사 결과는 정밀 발달평가 및 임상적 판단과 함께 종합적으로 해석되어야 한다는 점을 고려하였다.

III 평가결과

1. 문헌고찰

1.1 연구선정 개요

평가 주제와 관련된 연구를 찾기 위해 국내외 전자데이터베이스와 수기검색을 사용하여 총 895편의 연구를 확인하였다. 이 중 각 데이터베이스에서 중복 검색된 95편을 제외한 800편을 대상으로 제목과 초록을 검토하였고, 평가 주제와 연관 있는 96편에 대해 원문을 검토하였다. 이후 연구선택 기준에 따른 선택 과정을 거쳐 최종적으로 덴버발달검사에 대한 연구 총 15편을 선택하였다. 본 평가의 최종 연구선정 흐름도는 배제사유를 포함하여 〈그림 3.1〉에 제시하였다.

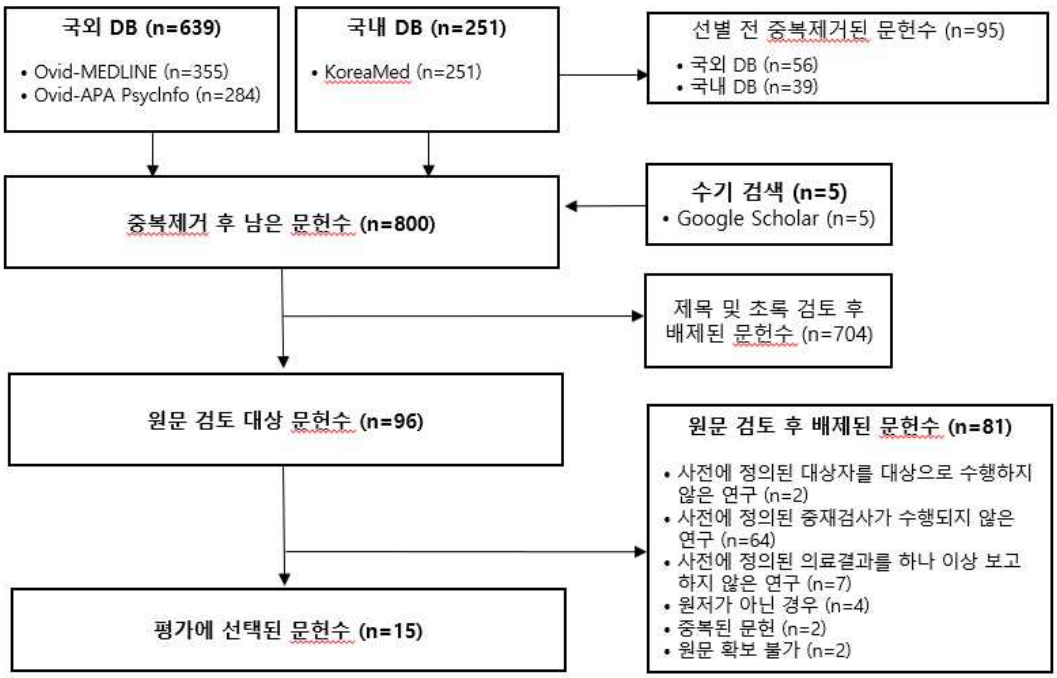

그림 3.1 연구선정 흐름도

1.2 선택연구 특성

본 평가에서는 덴버발달검사의 임상적 유용성과 관련된 총 15편의 문헌을 선택하였다. 선택연구는 평가목적에 따라 덴버발달검사의 타당성을 보고한 연구 8편, 선별검사로서의 정확성을 보고한 연구가 8편이 포함되었으며, 이 중 1편은 두 범주에 중복 포함되었다.

선택연구의 대상자는 크게 일반 아동 또는 정상 발달 아동, 발달지연 의심 또는 발달지연 위험요인을 가진 아동, 미숙아 또는 극소저체중 출생아와 같은 고위험군으로 구분되었다. 일반 아동 또는 정상 발달 아동을 대상으로 한 연구는 주로 Denver II의 개발 및 재표준화, 국가별·문화권별 규준 비교, 번역판 또는 지역 적응판의 표준화와 관련된 연구였다. 발달지연 의심 또는 위험 아동을 대상으로 한 연구는 임상 환경 또는 선별검사 환경에서 덴버발달검사의 선별 정확도와 도구 간 일치도를 평가하였다. 미숙아 또는 극소저체중 출생아를 대상으로 한 연구는 고위험군 추적관찰 과정에서 발달지연 또는 이후 발달문제를 선별·예측하는 데 덴버발달검사가 활용될 수 있는지를 평가하였다. 선택연구의 특성은 〈표 3.1〉과 같다.

한편, 한국판 덴버발달검사에 대한 연구 7편은 국내 표준화 및 실제 활용 현황을 별도로 평가하기 위하여 본 장에서는 제외하고, 영유아발달검사(한국판 덴버발달검사) 보고서에서 별도로 다루었다.

표 3.1 선택연구의 특성

연번	제1저자 (출판연도)	연구국가	연구설계	대상자	대상자 특성
1	Frankenburg (1992)	미국	개발·표준화 연구	0-6세 아동 2,096명	일반 아동을 대상으로 한 Denver II 개발 및 재표준화 표본
2	Glascoe (1993)	미국	진단 정확도 연구	3-72개월 아동 104명	보육시설 이용 아동을 대상으로 한 일반 선별 환경연구
3	Lee (1996)	한국	관찰연구/ 규준 비교연구	정상아동 1,018명	서울지역 정상 영유아를 대상으로 한 국내 규준 비교연구
4	Jeon (1997)	한국	진단 정확도 연구	발달지연 의심아동 244명	발달지연을 주소로 내원한 임상 의심군
5	Shin (2002)	한국	비교연구/ 규준 비교 연구	0-6.5세 정상아동 1,054명	국내 정상 아동을 대상으로 한 문화간 규준 비교연구
6	Sondaar (2008)	네덜란드	진단 정확도 평가 연구	조산아 또는 극소 저체중 출생아 368명	인지발달 문제 위험이 있는 미숙아 고위험군
7	Shahshahani (2010)	이란	도구 타당도 검증 연구	0-6세 정상아동 221명	아동건강클리닉 방문 정상아동 대상 번역판 검증연구
8	Ga(2011)	한국	후향적 단면연구	4-60개월 아동 111명	K-ASQ와 Denver II를 모두 시행한 발달평가 대상아동
9	Wijedasa (2012)	스리랑카	단면 규준화 연구	0-80개월 정상아동 4,251명	아동건강클리닉 이용 정상아동 대상지역 규준화 연구
10	Srinithiwat (2014)	태국	단면연구	18, 24, 30개월 영유아 90명	정상발달 아동과 발달지연 의심아동을 포함한 도구 비교연구
11	De-Andrés -Beltrán (2015)	스페인	도구 타당도 검증 연구	9개월-6세 아동 200명	일반아동과 발달지연 아동을 함께 포함한 판별 타당도 평가연구
12	Abessa (2016)	에티오피아	지역 표준화 연구	생후 4일-70개월 영유아 1,597명	비서구 저소득 국가 일반 영유아 대상 문화적 적응 및 지역 표준화 연구
13	Jeong	한국	후향적 진단 타당도	미숙아 69명	신경발달장애가 의심되는 고위험

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	대상자 특성
	(2017)		검증 연구		미숙아
14	Santos (2022)	브라질	도구 타당도 검증 연구	0-72개월 아동 254명	조산, 저체중, 낮은 사회경제적 수준 등 발달지연 위험요인을 가진 아동
15	Christovao (2023)	브라질	후향적 코호트 연구	미숙아 121명	34주 미만 또는 출생 체중 1,500g 이하로 출생한 장기추적 미숙아

1.3 검사도구의 타당성

Denver II의 측정학적 특성 및 국내외 적용 가능성을 확인하기 위하여 도구 개발 및 재표준화 연구 1편과 국가별 표준화·적응 연구 및 신뢰도·타당도 평가 연구 8편 등 총 9편의 연구를 검토하였다. 검토된 연구에서는 문항 구성 및 규준 개발 과정, 검사자 간 신뢰도와 검사-재검사 신뢰도, 내부 일관성, 타 도구 및 임상평가와의 일치도, 국가별 규준 차이 등이 보고되었다 〈표3.2〉.

Frankenburg et al. (1992)는 기존 Denver Developmental Screening Test(DDST)의 규준 노후화, 문항 구성의 제한, 임상 적용 과정에서 제기된 개선 요구를 반영하여 Denver II를 개발하고 재표준화한 과정을 제시하였다. 연구는 0-6세 아동 2,096명을 대상으로 수행되었으며, 성별, 인종, 모친 학력, 거주 지역을 층화 변수로 구성한 할당표본 설계를 적용하였다. 재표준화 과정에서 잠정 문항 336개가 구성되었고, 문항별 통과율 곡선을 산출하기 위해 평균 540회 반복 시행을 통해 자료를 축적하였다. 문항별로 25·50·75·90 백분위 통과연령을 산출하였으며, 로지스틱 회귀 분석을 통해 연령에 따른 통과 확률을 모델링하고 집단 간 규준 차이를 검토하였다. 또한 Composite percentile(가중 평균 방식)을 활용하여 인구집단 특성에 따른 규준 변동을 종합적으로 평가하였다.

문항 선정 과정에서는 거부율, 무응답률, 검사자 수용성(검사 현장에서의 시행 가능성), 규준 차이의 크기, 검사자 간 신뢰도, 검사-재검사 신뢰도 등 복수 기준을 적용하여 최종 125개 문항이 확정되었다. 항목 구성 개정에서 언어 영역 문항 수는 기존 DDST 대비 86% 증가한 것으로 제시되었고, 발음 명료도 관련 항목(2개), 행동평가 척도(순응도, 두려움, 주의집중 등), 새로운 연령 눈금, 판정 체계의 보완('주의(Caution)' 범주)이 포함되었다. 재표준화 절차의 신뢰도 평가에서는 최종 문항의 검사자 간 일치율이 대부분 90% 이상으로 보고되었으며, Cohen's κ 계수 기준으로 전체 문항이 K ≥0.75 수준의 검사자 간 신뢰도를 보였다고 제시되었다. 또한 7-10일 간격의 검사-재검사에서도 다수 문항에서 K ≥0.75 수준이 보고되어 시간 간격을 둔 측정의 안정성 자료가 함께 제시되었다.

Lee(1996)는 서울 지역 영유아 1,018명(남아 54.3%, 여아 45.7%)을 대상으로 1994년 12월부터 1995년 2월까지 Denver II를 적용하여 원판 규준과의 비교 자료를 제시하였다. 연구에서는 문화적 요인으로 인해 국내 적용이 어려운 9개 항목을 제외하고 116개 항목을 분석 대상으로 하였다. 검사자 간 신뢰도(inter-rater reliability)는 95.6-100%로 보고되었고, 평균 일치율은 98.2%로 제시되었다. 규준 비교는 문항별 90 백분위 통과연령을 산출한 뒤 원판 규준과 비교하는 방식으로 수행되었으며, 전체 31개 항목(26.7%)에서 20% 이상의 연령 차이가 보고되었다. 해당 비교 결과는

서울 지역 아동이 특정 항목에서는 더 느리고(16개 항목), 일부 항목에서는 더 빠른(15개 항목) 통과연령 분포를 보이는 양상으로 제시되었다. 연구에서는 외국 규준을 그대로 적용할 경우 판정 결과에 영향을 줄 수 있음을 기술하고, 국내 아동 특성을 반영한 표준화 필요성을 근거와 함께 제시하였다.

Shahshahani et al. (2010)는 Tehran 지역의 0-6세 소아 221명을 대상으로 Persian version of Denver II의 측정학적 특성을 평가하였다. 신뢰도 측면에서 검사자 내 신뢰도는 ICC 0.90-0.95, 검사자 간 신뢰도는 ICC 0.86-0.96으로 보고되었다. 내부 일관성은 Kuder-Richardson coefficient 0.61-0.74 범위로 제시되었다. 타 도구 또는 임상 평가와의 일치도 측면에서는, Denver II 결과와 Ages and Stages Questionnaires(ASQ) 간 발달지연 판정의 일치계수(kappa)가 0.21로 보고되었으며, 소아과 전문의의 임상평가와의 일치계수는 0.44로 제시되었다. 본 연구에서는 도구 번역·적용 맥락과 함께 측정 지표(ICC, 내부 일관성, 도구 간 판정 일치도)가 수치 형태로 제시되었다.

Wijedasa(2012)는 스리랑카 0-80개월 아동 4,251명을 대상으로 Denver II의 현지 적응 및 지역 규준(reference norms) 구축을 수행하였다. 연구에서는 연령별로 검사 항목의 통과연령 분포를 산출하고, 90번째 백분위수 기준을 적용하여 문항별 기준치를 제시하였다. 또한 산출된 지역 규준을 미국 원판 및 싱가포르 규준과 비교하여, 일부 항목에서 발달 성취 연령의 차이가 보고되었다. 결과는 문항 수준의 통과연령 차이를 제시하는 방식으로 정리되었다.

De-Andrés-Beltrán et al. (2015)는 9개월-6세 아동 200명을 대상으로 Spanish Version of Denver II의 심리측정학적 특성을 분석하였다. 신뢰도 지표로 검사자 내 신뢰도(Cohen's kappa 0.662-1), 검사자 간 신뢰도(0.886-1)가 제시되었고, 내부 일관성은 Kuder-Richardson coefficient 0.976으로 보고되었다. 본 연구는 스페인어판 도구에 대해 신뢰도와 내부 일관성 지표를 수치로 제시하였다.

Abessa et al. (2016)는 에티오피아 Jimma 지역의 생후 4일-70.6개월 소아 1,597명을 대상으로 Denver II를 현지 언어(Amharic, Afan Oromo)로 번역·적응한 뒤 표준화하였다. 원래 125개 항목 중 36개 항목(28.8%)이 문화·언어적 이유로 수정되었고, 통과연령(milestone attainment age)은 42개 항목(약 33.6%)에서 원판 규준과 차이가 보고되었다. 신뢰도 측면에서 검사자 간 신뢰도는 123개 항목(98%)에서 보고되었고, 검사-재검사 신뢰도는 119개 항목(91%)에서 substantial to excellent 수준으로 제시되었다. 본 연구는 번역·문항 수정 규모와 문항 수준의 통과연령 차이, 그리고 항목 단위 신뢰도 평가 결과를 함께 제시하였다.

Santos et al. (2022)는 0-72개월 발달지연 위험군 254명을 대상으로 Brazilian version of Denver II의 심리측정학적 특성을 평가하였다. 신뢰도 지표로 검사자 내·검사자 간 신뢰도가 ICC 0.97-0.99 범위로 보고되었다. 동시타당도는 ASQ-3BR (Brazilian version of Ages & Stages Questionnaires-3)과의 비교를 통해 Spearman 상관계수로 제시되었으며, 전체 연령군에서 0.13-0.91 범위였다. 연령군별로는 1-12개월에서 0.13-0.69 범위, 13-60개월에서 0.67-0.91 범위로 제시되었다. 본 연구에서는 연령군별 상관계수 범위를 구분하여 제시하였다.

Shin et al. (2002)는 Denver II의 국내 적용과 관련하여, 서울·수도권·충남 지역의 0-6.5세 아동

1,054명(남아 537명, 여아 517명)을 대상으로 발달 항목별 통과연령 분포를 산출하고 원판 규준과 비교하였다. 자료 수집은 보건소, 어린이집, 소아과 외래 등 다양한 현장 기반 환경에서 수행되었으며, 연령군 할당표집을 통해 연령 분포를 구성하였다. 분석에는 로지스틱 회귀분석이 적용되었고, 문항별 25·50·75·90 백분위 통과연령이 산출되었다. 검사자 간 신뢰도는 Kappa ≥0.70 수준에서 96개 문항에서 일치가 확인된 것으로 보고되었다. 규준 비교 결과, 90% 통과연령 기준으로 45개 항목(36.6%)에서 미국 규준과 차이가 보고되었고, 그 중 일부 항목(9개)은 국내 아동이 더 이른 통과연령을 보인 것으로 제시되었다. 50% 통과연령 기준에서는 38개 항목에서 차이가 보고되었으며, 일부 항목(16개)에서 국내 아동이 더 이른 통과연령을 보인 것으로 제시되었다. 연구는 문항별 규준 차이의 규모와 방향을 제시하는 형태로 결과를 정리하였다.

종합하면, Denver II는 개발 및 재표준화 과정에서 체계적인 문항 선정 절차와 규준 개발 과정을 거쳐 구축되었으며, 다수의 연구에서 높은 수준의 검사자 간 신뢰도와 검사-재검사 신뢰도가 보고되었다. 또한 국가별 표준화 연구에서는 문화·언어적 특성에 따라 일부 문항의 통과연령 및 규준 차이가 확인되어 지역별 적응과 표준화의 필요성이 제시되었으나, 전반적으로 각국에서 적용 가능한 발달선별도구로 활용되고 있었다. 타당도 측면에서는 임상평가 및 다른 발달선별검사와의 비교를 통해 일정 수준의 일치도와 상관성이 보고되었으며, 국내 연구에서도 미국 원판 규준과의 차이를 확인하고 한국 아동의 특성을 반영한 규준 개발의 필요성이 제시되었다.

표 3.2 덴버발달검사의 타당성

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
1	Frankenburg (1992)	미국	도구 개발 및 재표준화 연구	0-6세 아동 2,096명 (Denver County 1,039명 + Colorado state 1,057명)	Denver II	N/A	N/A	- 잠재 문항 336개를 평균 540회 시행하여 최종 125개 문항으로 선정 - 백분위(25·50·75·90) 통과연령 규준을 로지스틱 회귀모델로 산출 - 검사자 간 신뢰도(inter-rater agreement): 모두 Kappa ≥ 0.75(Excellent) - 7~10일 후 검사-재검사 신뢰도: 평균 0.90 (SD 0.12) - 언어 문항 86% 증가, parent-report 문항 20% 감소, 발음 intelligibility 2개 항목 추가 - caution/delay 기준 도입: delay(≥90% 통과연령 실패), caution(75-90% 범위 실패) - 행동평가 척도 및 training manuals/video 제공 - 규준은 성별, 인종, 지역, 부모교육수준별 subgroup 규준 포함
2	Lee (1996)	한국	관찰연구	병원에 내원한 1,018명의 정상 아동 - 배제: 미숙아, 쌍생아, 둔위 분만아, 입양아, 시청각 장애아, 중추 신경 장애, 구개 파열 및 다운 증후군, 급성 혹은 만성 질환 아동	Denver II (한국어)	N/A	N/A	- Denver II를 서울 지역 영유아에게 적용했을 때 미국 원판 규준과의 발달 연령 차이와 지역 적용 가능성을 평가 - 검사자 간 신뢰도(inter-rater reliability): 95.6-100%, 평균 98.2%로 매우 우수 - 발달 연령 차이 존재: 전체 116개 문항 중 31개 항목(26.7%)에서 20% 이상 차이 - 문화·환경 요인에 따른 규준 차이 가능성 제기: 발달 환경, 자극, 교육체계 등 요인 고려 필요 - 외국 규준 적용 시 발생 가능한 문제: 발달지연 아동을 정상으로 오판할 위험, 정상 아동을 과대 선별할 위험 - 한국 아동 발달특성을 반영한 표준화의 필요성을 제시한 최초의 근거 연구 - 이후 한국판 덴버발달검사 표준화 작업의 근거 기반 형성

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
3	Shin (2002)	한국	비교연구	0~6세(0~78개월) 정상 아동 1,054명 - 배제: 미숙아, 선천성 기형, 신경학적 이상, 감각장애, 만성질환	Denver II (한국어)	N/A	N/A	- Denver II의 문화·지역 간 규준 차이를 탐색한 규준 비교연구 - 총 125개 항목 중 45개 항목(36.6%)에서 90번째 백분위 통과연령 기준으로 미국 규준 대비 유의한 발달 차이 확인 - 이러한 차이는 양육방식, 사회문화적 환경, 언어체계의 차이에 기인한 것으로 해석됨 - 미국 규준을 그대로 적용할 경우 발달지연 판정의 과대·과소 선별 위험 존재, 발달지연 의심 아동이 정상으로 분류되거나, 정상 아동이 과다 선별되는 오류 가능성 있음 - 한국 아동 발달 특성을 반영한 재표준화 연구 필요성을 제언하며, 국가 차원의 표준화 및 검증 작업의 근거 연구로 활용될 수 있음을 제시함
4	Shahshahani (2010)	이란	도구 타당도 검증 연구	4개의 아동 건강 클리닉 방문한 0~6세 정상 아동 221명 - 배제: 명백한 발달지연 또는 장애 보유 아동	Denver II (이란어)	ASQ (4~60개월 대상)	일부 아동 (ASQ 적용불가군 및 10%)에 대해 소아발달 전문의 임상평가	- 페르시아어로 번역·적용된 Denver II의 신뢰도와 타당도를 평가하였음. 검사자 내 신뢰도(ICC=0.90–0.95, κ=0.87), 검사자 간 신뢰도 (ICC=0.86–0.96, κ=0.76)는 매우 우수하였으며, 내적 일관성 (Kuder–Richardson 0.61–0.74) 또한 양호하였음 - ASQ 대비 민감도 0.60, 특이도 0.69, 소아과 전문의 평가 대비 일치도 κ=0.44로 나타나, 이란판 Denver II는 전반적으로 신뢰할 수 있는 발달선별 도구로 평가되었음 - 다만 ASQ와의 일치도는 낮아(κ = 0.21), 추가적인 표준검사 기반 검증이 필요함을 제시하였음
5	Wijedasa (2012)	스리랑카	단면 규준화 연구	11개 구의 아동건강클리닉을 이용하는 0~80개월 정상 아동 4,251명 - 포함: 만삭 출생, 발달장애 진단 받은 적 없고, 만성질환이 없는 아동	Denver II (스리랑카어)	N/A	N/A	- 스리랑카 아동에게 DDST-II를 적용·표준화하여 자국 발달규준(DDST-SL)을 확립하고, 그 결과를 미국 DDST-II·싱가포르 DDST·스리랑카 CHDR (아동건강발달기록) 규준과 비교한 연구로 각 항목의 25·50·75·90 백분위 통과연령을 제시하였음 - DDST-SL 규준의 90th percentile 통과연령은 DDST-II 및 DDST-Singapore 규준과 유의한 차이를 보였음 - 국가별 문화·환경적 차이에 따른 발달속도의 편차를 반영한 현지 표준화의 필요성을 제시함

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
6	De-Andres-Beltran (2015)	스페인	도구 타당도 검증연구	9개월~6세 아동 200명 - 정상군: 일반 유치원 아동 100명 - 발달지연군: EIC (Early Intervention Centres) 등록 아동 100명 (소아과 전문의 확진)	Denver II (스페인어)	N/A	소아과 전문의 임상진단 (Psychomotor delay 존재 여부)	- 스페인어 버전의 Denver II의 신뢰성과 타당성 연구 - 민감도: 0.89, 특이도: 0.92, 양성예측도: 0.92, 음성예측도: 0.89로 발달지연 아동을 선별하는 정확도가 높았음 - 양성우도비: 11.1, 음성우도비: 0.12로 검사 결과의 임상적 유용성이 높았음 - 검사자 내 신뢰도는 카파값 0.662~1.00(평균 일치율 98%), 검사자 간 신뢰도는 카파값 0.886~1.00(평균 일치율 96%)로 모두 매우 높았음 - 내적 일관성은 전체 0.976으로, 모든 하위 영역에서 높은 일관성을 보였음 - 스페인어판 Denver II가 발달지연의 조기 선별에 적합하며, 우수한 타당도와 신뢰도를 가진 도구임
7	Abessa (2016)	에티오피아	지역 표준화 연구	생후 4일~70개월 사이 영유아 1,597명	Denver II (에티오피아어)	N/A	Denver II	- 에티오피아 Jimma 지역 아동을 대상으로 Denver II의 번역·적응판의 적용 가능성과 지역 규준 구축을 위한 연구 - 원판 125개 항목 중 36개 항목(28.8%)이 문화·언어적 차이로 수정되었고, 42개 항목(33.6%)에서 통과연령(milestone attainment age)이 미국 원판과 유의하게 차이가 나타남 - 검사자 간 신뢰도는 전체 123개 항목(98%)에서 substantial~excellent 수준으로 확인됨 - 재검사 신뢰도는 119개 항목(91%)에서 substantial~excellent 수준으로 보고됨 - 국가 및 문화적 특성에 따라 발달 규준과 항목 난이도가 달라질 수 있어, 일률적인 원판 규준 사용 시 판정 오류 가능성이 제기됨 - Denver II의 지역 표준화와 국가별 규준 개발의 필요성을 확인한 연구
8	Santos (2022)	브라질	도구 타당도 검증연구	발달지연 위험이 있는 0-72개월 아동 254명 - 포함: 생물학적·환경적 위험요인 (조산, 저체중, 임신합병증, 낮은 사회경제적 수준 등)을 가진 아동 - 제외: 확정된 신경학적·유전적 질환 (다운증후군, 뇌성마비 등)을 가진 아동	Denver II (브라질어)	ASQ-3 (브라질어)	N/A	- 검사자 내·간 신뢰도(ICC 0.97-0.99), 동시타당도(r 0.67-0.91), 민감도 평균 0.89, 특이도 평균 0.78 - 1~12개월에서는 상관 낮음(r 0.13-0.69), 13-60개월에서는 중등도-강한 상관(r 0.67-0.91) - Denver II는 ASQ-3BR이 평가하지 못하는 61-72개월 연령대에서도 적용 가능하며, 특히 13개월 이상의 아동에서 발달지연 선별의 타당도와 정확도가 높게 확인되었음 - Denver II와 ASQ-3BR의 정성적 결과를 통해, Denver II가 연령 증가에 따라 민감도와 특이도의 좋은 비율 지표를 제공한다는 것을 확인했음

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과

NA, Not applicable ;
KU-Binet, Korea University-Binet Intelligence Test ;
DDST-II, Denver Developmental Screening Test-II ;
DCD, Developmental Coordination Disorder ;
ASQ-3, Ages & Stages Questionnaires-3 ; K-ASQ, Korean Ages & Stages Questionnaires ;
Bayley-II (Bayley Scales of Infant Development-II) ; Bayley-III (Bayley Scales of Infant and Toddler Development-III) ;
Movement ABC, Movement Assessment Battery for Children ;
TST, Taal Screening Test ; RAKIT. Revised Amsterdam Child Intelligence Test ;
ICC, Intraclass Correlation Coefficient ;
MDI, Mental Developmental Index ; PDI, Psychomotor Developmental Index.

1.4 선별검사로서의 정확도

Denver II의 발달지연 선별검사로서의 정확도를 확인하기 위하여 총 8편의 연구를 검토하였다. 연구들은 일반 영유아, 발달지연 의심 아동, 미숙아 등 다양한 대상군을 포함하였으며, 발달전문평가, Bayley Scales of Infant Development(BSID), 지능검사, 임상진단, Ages and Stages Questionnaire(ASQ) 등 다양한 참고표준 또는 비교도구와 Denver II 결과를 비교하였다. 검토된 연구에서는 민감도, 특이도, 양성예측도, 음성예측도, 일치도(kappa), 예측타당도 등의 지표를 통해 Denver II의 발달지연 선별 성능을 평가하였다 〈표3.3〉.

Glascoe et al.(1993)는 임상 환경에서 Denver II의 선별 정확도를 평가하였다. 연구는 3-72개월 아동 104명을 대상으로 수행되었으며, Denver II 결과는 독립적으로 실시된 발달 전문 검진(reference diagnostic battery)과 비교되었다. 참고표준에는 표준화된 지능검사, 언어 및 학습능력 평가, Vineland Adaptive Behavior Scales, 소아과 전문의 임상 진단 결과가 포함되었다. 연구는 Denver II의 판정 기준에 따라 정확도 지표가 달라짐을 수치로 제시하였다. 기본 판정 기준(default scoring; questionable 항목을 pass 처리) 적용 시 민감도 56%, 특이도 80%, 양성예측도 37%로 보고되었다. 반면 questionable 항목을 fail 범주에 포함하는 수정 판정 기준(revised scoring) 적용 시 민감도 83%, 특이도 43%, 양성예측도 23%로 보고되었다. 또한 영역별로는 언어 영역에서 판정 일치율이 상대적으로 높게 나타난 반면, 개인-사회성, 소근육-적응, 대운동 영역에서는 판정 결과 변동이 관찰된 것으로 기술되었다.

Jeon et al. (1997)의 연구는 대학병원 재활의학과에 발달지연을 주소로 내원한 소아 244명을 대상으로 Denver II의 발달지연 선별 성능을 분석하였다. 전체 민감도는 0.85, 특이도는 0.59로 확인되었고, 특히 언어 발달지연군에서는 민감도 0.90, 특이도 0.27로 나타나 발달 위험군 탐지에는 높은 민감도를 보이나 위양성률이 높아 단독 검사로서는 제한이 있음을 제시하였다. 연구진은 Denver II를 진단 도구가 아닌 선별 도구로 활용해야 하며, BSID-II (Bayley Scales of Infant Development-II) 등 정밀평가 도구 및 임상적 판단과의 병행이 필수적이라고 강조하였다.

Sondaar et al. (2008)는 재태연령 32주 미만 또는 출생체중 1,500g 미만으로 출생한 고위험 미숙아 368명을 대상으로 Denver II의 발달지연 선별 성능을 평가하였다. 연구에서는 발달지연 여부 판단의 참고표준으로 Revised Amsterdam Child Intelligence Test (RAKIT) 지능검사를 사용하였으며, 지능지수(IQ) 85 미만을 발달지연 기준으로 정의하였다. Denver II는 교정연령에 맞추어 소아과 전문의에 의해 시행되었으며, RAKIT 결과와의 비교를 통해 선별 정확도가 분석되었다. 결과에서 민감도 0.72, 특이도 0.79, 양성예측도 0.40, 음성예측도 0.93이 보고되었으며, 이는 Denver II가 발달 문제의 가능성이 낮은 대상을 배제하는 데 유용한 선별검사적 특성을 보였음을 의미한다. 연구에서는 고위험 미숙아 추적관리 과정에서 발달 전문평가(정밀검사) 의뢰 여부를 결정하는 초기 단계 선별 도구로 활용될 수 있는 근거를 제시하였다.

Shahshahani et al. (2010)는 ASQ를 비교 도구로 사용하여 선별 정확도 지표를 함께 제시하였다. ASQ와의 비교에서 Denver II의 민감도는 60%, 특이도는 69%로 보고되었다.

Ga et al.(2011)는 발달지연 선별검사로서 K-ASQ (Korean Ages and Stages Questionnaires)의 유용성을 평가하기 위해 Denver II 및 참고표준검사(Capute scale, Alberta Infant Motor Scale [AIMS])와의 동시타당도를 비교하였다. 연구는 4-60개월 소아 226명을 대상으로 하였으며 이 중 111명은 Denver II를 동일 시점에 병행 시행하였다. 발달지연 위험군 분류에서 K-ASQ는 73명, Denver II는 95명을 위험군으로 확인하였고, 두 검사 간 일치도는 kappa=0.442 수준으로 보고되었다. 참고표준검사(Capute scale 또는 AIMS 기준) 대비 K-ASQ 민감도는 76.3-90.2%, 특이도는 62.5-76.5%, 동일 기준에서 Denver II는 민감도 92.0-97.0%, 특이도 25.0-38.9%였다.

Srinithiwat et al. (2014)는 태국 방콕 지역 Queen Sirikit National Institute of Child Health에서 18, 24, 30개월 소아 90명을 대상으로 ASQ-3 Thai와 Denver II를 비교하여 선별 결과를 제시하였다. 두 도구 간 전체 일치율은 연령군별로 kappa 0.338-0.606 범위로 보고되었다. ASQ-3 Thai 기준 위험군과 Denver II 기준 위험군 간 비교에서 민감도는 54.5-71.4%, 특이도는 78.6-90.9%로 보고되었다. 본 연구는 특정 월령군에서 도구 간 비교를 통해 산출된 정확도 지표를 제시하였다.

Jeong et al. (2017)는 신경발달장애가 의심되는 미숙아 69명을 대상으로 Denver II와 Bayley-III의 발달지연 선별 결과를 비교하였다. 기존 판정 기준 적용 시 Denver II의 발달지연 선별률은 50.7%로 보고되었고, Bayley-III의 선별률은 31.88%로 제시되었다. 판정 기준을 '주의(Caution)'이 1개 이상인 경우를 이상(abnormal)으로 간주하는 기준'으로 적용할 경우 Denver II의 선별률은 69.6%로 보고되었으며, 교정연령 13-24개월군에서는 Denver II의 선별률이 100%로 제시되었다. 두 검사 간 발달지연 판정의 일치도는 kappa 0.048(p=0.664)로 보고되었다. 본 연구는 동일 표본에서 판정 기준 변화에 따른 선별률 차이와 도구 간 일치도 지표를 제시하였다.

Christovao et al.(2023)는 미숙아 121명을 대상으로 1~5세 시점의 Denver II 결과와 7세 시점의 운동발달평가(Movement Assessment Battery for Children-Second Edition, MABC-2)을 활용하여 운동발달을 평가하였다. 연구 결과, Denver II에서 suspect로 분류된 아동은 정상군에 비해 학령기 운동발달장애(Developmental Coordination Disorder, DCD) 위험이 3.45배에서 7.40배까지 높은 것으로 보고되었다($p<0.01$). 반복된 검사에서 suspect 판정 횟수가 증가할수록 위험도가 상승하는 경향이 확인되었다. Denver II 결과와 7세 시점 발달평가 결과 간에는 통계적으로 유의한 관련성이 보고되었다.

종합하면, Denver II는 대부분의 연구에서 발달지연 위험군을 선별하는 데 비교적 높은 민감도를 보였으나, 특이도는 연구 및 적용 기준에 따라 다소 낮거나 변동이 있는 것으로 보고되었다. 특히 발달지연 가능성이 있는 아동을 놓치지 않는 선별검사로서의 특성은 확인되었으나, 위양성 결과가 발생할 수 있어 단독으로 진단에 활용하기보다는 정밀 발달평가 및 임상적 판단과 함께 사용하는 것이 필요하다고 제시되었다. 또한 일부 연구에서는 향후 발달문제 및 운동발달장애 위험을 예측하는 데 유의한 관련성이 보고되어 선별도구로서의 활용 가능성을 제시하였다. 전반적으로 Denver II는 발달지연 위험군을 조기에 확인하여 추가 평가가 필요한 대상을 선별하기 위한 목적으로 활용 가능한 것으로 보고되었다.

표 3.3 덴버발달검사의 선별검사로서의 정확도

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
1	Glascoe (1993)	미국	진단 정확도 연구	5개의 보육시설의 3~72개월 아동 104명	Denver II	N/A	(1) cognitive test (Bayley / K-ABC /Stanford-Binet) (2) Vineland Adaptive Behavior (3) K-ABC Achievement	- Denver II는 발달검사로서 높은 민감도(83%)를 보였으나, 특이도(43%)와 양성예측도(23~37%)는 낮았음 - 실제로 선별된 아동의 약 세 명 중 두 명은 추가 진단에서 정상으로 판정될 가능성이 높아, 불필요한 과잉 의뢰가 빈번히 발생함 - 참고표준(reference standard)은 단일검사가 아니라 실제 임상진단 절차를 반영한 복합 진단배터리(diagnostic battery)로 구성 - Denver II는 발달지연을 조기에 포착하는 능력은 양호하나, 정상 아동을 과잉 선별하는 경향이 뚜렷하여 민감도는 높지만 특이도는 낮은 선별검사로 평가
2	Jeon (1997)	한국	진단 정확도 연구	발달지연이 의심되어 발달클리닉에 내원한 아동 244명	Denver II (한국어)	N/A	임상적·신경학적 진단	- 그룹별 민감도/특이도 · 운동지연군 0.88 / 0.41 · 언어지연군 0.90 / 0.27 · 기타군 0.81 / 0.77 · 전체 0.85 / 0.59 - 발달지연 위험이 있는 아동을 식별하는 Denver II는 민감도가 우수하지만, 특히 언어 지연 위험이 있는 아동을 식별하는 데 있어 특이성이 낮음 - 발달지연이 있는 한국 어린이에게 사용할 수 있도록 표준화되고 수정되어야 함
3	Sondaar (2008)	네덜란드	진단 정확도 평가 연구	5세 미만 인지장애 미진단 조산아(임신 32주 미만 또는 1,500g 미만 극소저체중) 368명 - 배제: 중증 장애 또는 이미 인지장애 진단된 아동	소아청소년과 전문의의 발달평가 (신체·신경학적 평가 + 부모보고 + DDST + 언어검사(TST)	N/A	RAKIT 지능검사 (Dutch intelligence test)	- DDST 단독 결과: 민감도: 0.72, 특이도: 0.76, 음성예측도: 0.93, 음성우도비: 0.35 - 소아청소년과 전문의 임상판단(종합평가) 민감도: 0.42로 낮아, 실제 인지장애 아동의 60% 이상을 인식하지 못함 - 학교 수행 정보 함께 고려했을 때 민감도: 0.79, 음성예측: 0.94 (IQ와 높은 일치) - DDST는 소아과 발달평가 도구 중 높은 민감도를 보여 조기 인지장애 식별에 유용한 선별검사로 평가됨. 다만, 단독검사로는 일부 인지저하 아동 놓칠 수 있어, 필요시 심리학적 정신 지능검사로의 연계가 권장됨

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
4	Shahshahani (2010)	이란	도구 타당도 검증 연구	4개의 아동 건강 클리닉 방문한 0~6세 정상 아동 221명 - 배제: 명백한 발달지연 또는 장애 보유 아동	Denver II (이란어)	ASQ (4~60개월 대상)	일부 아동 (ASQ 적용불가군 및 10%)에 대해 소아발달 전문의 임상평가	- 페르시아어로 번역·적용된 Denver II의 신뢰도와 타당도를 평가하였음. 검사자 내 신뢰도(ICC=0.90-0.95, κ=0.87), 검사자 간 신뢰도 (ICC=0.86-0.96, κ=0.76)는 매우 우수하였으며, 내적 일관성 (Kuder-Richardson 0.61-0.74) 또한 양호하였음 - ASQ 대비 민감도 0.60, 특이도 0.69, 소아과 전문의 평가 대비 일치도 κ=0.44로 나타나, 이란어 Denver II는 전반적으로 신뢰할 수 있는 발달선별 도구로 평가되었음
5	Ga (2011)	한국	후향적 단면연구	생후 4개월에서 60개월 사이의 아동 중 K-ASQ와 DDST II를 모두 실시한 아동 111명	Denver II (한국어)	K-ASQ	Capute Scale 및 Alberta Infant Motor Scale(AIMS)	- K-ASQ와 Denver II의 일치도 (Kappa=0.442, p〈0.05) → 보통 수준의 동시타당도 확인 - 발달지연 위험군 분류율: K-ASQ 65.8%, DDST-II 85.6% → 덴버가 더 많은 아동을 위험군으로 분류 - Capute Scale(DQ〈75) 기준: Denver II 민감도 97.0%, 특이도 25.0%, K-ASQ 민감도 90.2%, 특이도 62.5% - AIMS 기준: Denver II 민감도 92.0%, 특이도 37.9%, K-ASQ 민감도 80.0%, 특이도 76.5% - Denver II는 높은 민감도를 지닌 표준 선별검사이며, K-ASQ는 그 보완적 도구로서 효율적이고 실용적인 가치가 있다고 제시함
6	Srinithiwat (2014)	태국	단면연구	총 90명 (18개월, 24개월, 30개월 영유아 각 30명씩) 각 연령군별로 정상 발달 아동 15명, 발달지연 의심 아동 15명 포함	Denver II	ASQ-3 (태국어)	N/A	- ASQ-3 Thai는 Denver II와 비교 시 동시타당도(Kappa 0.34~0.61) 수준으로, 보통 정도의 일치도를 보였음 - ASQ-3 Thai는 특이도가 높고 부모 참여형 도구로서 접근성이 우수하나, 민감도가 낮아 단독 사용 시 발달지연 아동의 일부를 놓칠 가능성이 존재하므로, 임상 현장에서는 Denver II 등 타 표준화 도구와 병행 사용이 권장됨 - DDST-II가 적절한 심리측정학적 특성을 가지고 있어 광범위한 발달 선별 도구로 널리 사용되고 있음. 신뢰도 0.99±0.016, 민감도 0.56~0.8, 특이도 0.43~0.80 (* 비고: ASQ-3 Thai가 중재도구이며, Denver II는 비교·검증 목적의 참고표준으로 활용된 연구)

연번	제1저자 (출판연도)	연구 국가	연구설계	대상자	중재	비교중재	참고표준	주요결과
7	Jeong (2017)	한국	후향적 진단 타당도 검증연구	2010.6.~ 2013.8. 소아재활의학과를 방문했거나 입원한 37주 미만의 미숙아 69명 - 포함: 신경학적 검사에서 이상이 있고, Bayley-III와 Denver II를 모두 시행한 환아	Denver II (한국어)	Bayley III	신경학적 검사 (muscle tone, primitive reflex, Vojta reaction, cranial nerve, muscle power, deep tendon reflex, sensory function, and general movement)	- 고위험 미숙아에서 Bayley-III 및 Denver II의 발달지연 탐지율(detection rate) 비교 연구 - 탐지율: Denver II 50.7% (35/69) vs Bayley-III 31.9% (22/69) - Bayley-III가 교정연령 12개월 미만 영아군에서 인지 및 운동 영역의 발달지연을 더 잘 탐지하였으며, Denver II는 13-24개월군에서 언어 및 운동 영역의 발달지연을 100% 수준으로 검출하였음 - 두 검사 간 전체 일치도 κ=0.048로 낮아, 하나의 검사 결과로 다른 검사의 발달 판정을 대체할 수 없으며, 평가 목적과 연령에 따라 상보적으로 병행 활용하는 것이 적절함
8	Christovao (2023)	브라질	후향적 코호트 (2005.1 ~2019.12.)	34주 미만 또는 출생체중 ≤1500g 미숙아 121명 - 포함: 7세까지 ACRIAR 프로그램에서 추적 관찰된 아동 - 배제: 유전증후군, 선천성 기형, 진단된 신경학적 질환이 있는 아동, 7세 평가 시점에 추적 관찰에 참여하지 않은 아동	Denver II	N/A	Movement Assessment Battery for Children, 2nd Edition (MABC-2)	- 조산아의 DCD 위험을 규명하기 위한 Denver II의 예측 타당성 연구 - 생후 1~5세 시점에 시행된 Denver II의 검사 결과를 분석한 결과, 특히 2~4세 시점의 '의심' 판정이 7세 시점의 DCD 위험을 유의하게 예측함이 확인되었음 (민감도 66-70%, 특이도 51-55%) - 1~5세 사이에 '의심' 판정이 반복적으로 나타난 아동일수록 DCD 발생 가능성이 증가 - Denver II는 미숙아에서 DCD 위험을 조기에 예측할 수 있는 보조적 선별 도구로서의 예측타당도를 보였음

NA, Not applicable ;
KU-Binet, Korea University-Binet Intelligence Test ;
DDST-II, Denver Developmental Screening Test-II ;
DCD, Developmental Coordination Disorder ;
ASQ-3, Ages & Stages Questionnaires-3 ; K-ASQ, Korean Ages & Stages Questionnaires ;
Bayley-II (Bayley Scales of Infant Development-II) ; Bayley-III (Bayley Scales of Infant and Toddler Development-III) ;
Movement ABC, Movement Assessment Battery for Children ;
TST, Taal Screening Test ; RAKIT. Revised Amsterdam Child Intelligence Test ;
ICC, Intraclass Correlation Coefficient ;
MDI, Mental Developmental Index ; PDI, Psychomotor Developmental Index.

2. 전문가 의견

본 평가에서는 문헌고찰과 소위원회 논의만으로 확인하기 어려운 덴버발달검사의 국내 활용 맥락을 보완적으로 검토하기 위하여 전문가 자문을 실시하였다. 자문에서는 덴버발달검사의 검사 성격, 국내외 활용 동향, 지역사회 기반 활용 경험, K-DST 도입 이후의 역할 등이 검토되었다.

전문가 자문 결과, 덴버발달검사는 특정 발달 문제의 심각성을 판단하거나 이상 여부를 진단하기 위한 도구라기보다, 아동 발달의 전반적 흐름을 확인하고 발달지연 가능성이 있는 아동을 선별하기 위한 도구로 보는 것이 적절하다는 의견이 제시되었다. 즉, 덴버발달검사는 베일리 발달검사, 웩슬러 지능검사, CARS 등과 같은 정밀 진단·평가 도구와 목적이 다르며, 검사 결과는 발달 상태를 확정적으로 판단하는 근거가 아니라 정밀평가 의뢰, 추적관찰 또는 중재 연계 필요성을 판단하기 위한 참고자료로 해석할 필요가 있다.

덴버발달검사는 생애 초기 아동에게도 적용 가능하고, 검사자가 아동의 행동과 발달 상태를 직접 관찰하면서 시행할 수 있다는 특성이 있다. 이러한 특성으로 인해 생애초기 건강관리사업, 서울아기 건강 첫걸음 사업 등 일부 가정방문사업에서는 간호사, 사회복지사 등 비의료 인력이 별도의 훈련을 받은 후 덴버발달검사를 활용한 사례가 공유되었다.

국외 활용 동향과 관련해서는 덴버발달검사가 최근 주요 국외 가이드라인 및 일부 소아과학 교과서에서 주요 도구로 언급되지 않는 점이 확인되었다. 덴버발달검사는 개발된 지 오래된 검사이며, 민감도와 특이도 측면에서 다양한 대체 도구가 존재함에 따라 국외에서는 사용이 감소하는 추세로 이해된다. 다만 이는 덴버발달검사가 진단용 표준검사로서의 위치를 갖지 않는다는 점과 관련되며, 선별도구로서의 활용 가능성은 진단도구와 구분하여 해석할 필요가 있다.

K-DST 도입 이후 덴버발달검사의 역할과 관련해서는, 두 도구의 평가 방식 차이가 검토되었다. K-DST는 부모보고식 검사인 반면, 덴버발달검사는 훈련된 검사자가 아동을 직접 관찰하여 실시하는 검사이다. 두 도구는 모두 발달지연 선별을 목적으로 하지만 평가 방식이 다르므로, 상황에 따라 병행 또는 선택적으로 활용할 경우 선별정확도를 보완하는 데 도움이 될 수 있다. 또한 단일 검사도구만 사용할 경우 변별력이 제한될 수 있으므로, 다양한 선별도구를 임상 및 지역사회 상황에 따라 선택적으로 활용할 수 있는 체계를 유지할 필요가 있다는 의견이 제시되었다.

종합하면, 전문가 자문에서는 덴버발달검사를 진단도구가 아니라 발달지연 가능성을 조기에 확인하기 위한 선별도구로 이해하는 것이 적절하다고 보았다. K-DST 도입 이후에도 부모보고식 선별검사와 관찰 기반 선별검사의 역할 차이를 고려할 때, 덴버발달검사는 제한적·보완적으로 활용될 수 있다. 다만 국외 주요 가이드라인 및 일부 교과서에서 덴버발달검사가 주요 도구로 언급되지 않는 점과 다양한 대체 선별도구가 활용되고 있는 상황을 고려할 때, 국내 임상적 유용성을 판단할 때에는 선별도구로서의 역할과 실제 운영 맥락을 함께 고려할 필요가 있다.

IV 결과 요약 및 결론

1. 평가결과 요약

본 의료기술재평가에서는 덴버발달검사의 국내 임상적 유용성을 평가하기 위하여 관련 연구 총 15편을 검토하였다. 또한 문헌고찰만으로 확인하기 어려운 국내 활용 맥락, 국외 활용 동향, K-DST 도입 이후의 역할 등을 보완적으로 검토하기 위하여 전문가 자문을 실시하였다.

선택연구는 평가목적에 따라 검사도구로서의 타당도 및 신뢰도를 보고한 연구와 선별검사로서의 정확도를 보고한 연구로 구분하였다. 검사도구로서의 타당도 및 신뢰도를 보고한 연구는 총 8편으로, 덴버발달검사의 개발 및 재표준화, 국가별·문화권별 규준 비교, 번역·적응판의 신뢰도 및 타당도 검증과 관련된 연구가 포함되었다. 이들 연구에서는 덴버발달검사의 문항 구성과 연령별 통과연령 산출 과정, 검사자 간 신뢰도, 검사-재검사 신뢰도, 내부 일관성 및 다른 발달평가 도구와의 일치도 등이 보고되었으며, 전반적으로 신뢰도는 양호한 수준으로 제시되었다. 또한 일부 국가별 표준화 연구에서는 원판 규준과 문항별 통과연령 차이가 확인되어, 문화적·언어적 맥락을 고려한 해석과 지역별 표준화의 필요성이 제시되었다.

선별검사로서의 정확도를 보고한 연구는 총 8편으로, 일반 아동, 발달지연 의심 또는 위험 아동, 미숙아 등 고위험군을 대상으로 덴버발달검사의 민감도, 특이도, 예측도, 검사 간 일치도 및 선별률 등을 평가하였다. 이들 연구에서 덴버발달검사는 발달지연 위험군을 조기에 확인하는 선별도구로 활용 가능성이 보고되었으나, 연구대상자, 참고표준 및 판정 기준에 따라 정확도 지표에는 차이가 있었다. 특히 일부 연구에서는 민감도는 비교적 높게 보고된 반면 특이도는 낮거나 변동이 있어, 검사 결과만으로 발달지연 여부를 확정하기보다는 정밀 발달평가 필요성, 추적관찰 여부 및 임상적 판단을 함께 고려할 필요가 있는 것으로 제시되었다.

전문가 자문에서는 덴버발달검사의 검사 성격, 국내외 활용 동향, 지역사회 기반 활용 경험, K-DST 도입 이후의 역할 등이 검토되었다. 자문 결과, 덴버발달검사는 특정 발달 문제의 심각성을 판단하거나 이상 여부를 진단하기 위한 도구가 아니라, 아동 발달의 전반적 흐름을 확인하고 발달지연 가능성이 있는 아동을 선별하기 위한 도구로 보는 것이 적절하다는 의견이 제시되었다. 또한 덴버발달검사는 검사자가 아동의 행동과 발달 상태를 직접 관찰하면서 시행할 수 있어, 일부 가정방문사업 등 지역사회 기반 사업에서도 활용된 사례가 확인되었다. K-DST 도입 이후의 역할과 관련해서는, K-DST가 부모보고식 검사인 반면 덴버발달검사는 훈련된 검사자가 아동을 직접 관찰하여 실시하는 검사라는 차이가 확인되었다. 두 도구는 모두 발달지연 선별을 목적으로 하지만 평가 방식이 다르므로, 임상 및 지역사회 상황에 따라 병행 또는 선택적으로 활용할 경우

선별정확도를 보완하는 데 도움이 될 수 있다는 의견이 제시되었다.

종합하면, 문헌고찰에서는 덴버발달검사가 발달지연 위험군을 조기에 식별하기 위한 선별도구로 활용될 수 있음을 보여주는 제한적 근거가 확인되었다. 전문가 자문에서는 덴버발달검사의 실제 활용 맥락과 K-DST 이후의 보완적 역할이 검토되었다. 따라서 덴버발달검사는 발달지연 여부를 확정적으로 진단하기 위한 도구라기보다, 발달지연 가능성을 조기에 확인하고 후속 평가 또는 추적관찰 필요성을 판단하기 위한 선별도구로 이해하는 것이 적절하다.

2. 결론 및 제언

의료기술재평가 소위원회에서는 '덴버발달검사'에 대해 현재 평가결과에 근거하여 다음과 같이 제언하였다.

문헌검토 결과, 덴버발달검사의 타당도 및 선별정확도를 평가한 연구 1편과 임상적 활용 양상을 보고한 연구 6편이 확인되었다. 문헌상 덴버발달검사는 발달지연 위험군을 조기에 선별하고, 다양한 대상자군에서 발달의심군 분류, 고위험군 확인, 관련 요인 분석, 후속 평가 필요성 검토 등을 위한 참고자료로 활용되었다. 다만 선별정확도와 임상적 유용성에 대한 근거는 제한적이며, 검사 결과는 단독으로 해석하기보다 정밀평가 및 임상적 판단과 함께 종합적으로 해석할 필요가 있었다.

추가 전문가 자문에서는 덴버발달검사가 특정 발달 문제의 심각성이나 이상 여부를 진단하기 위한 도구가 아니라, 아동 발달의 전반적 흐름을 확인하고 발달지연 가능성이 있는 아동을 선별하기 위한 도구라는 점이 확인되었다. 또한 K-DST 도입 이후에도 부모 보고 기반 선별검사와 관찰 기반 선별검사의 역할 차이를 고려하여, 상황에 따라 병행 또는 선택적으로 활용될 수 있다는 의견이 제시되었다.

문헌검토와 전문가 자문 결과를 종합한 결과, 덴버발달검사가 장애 진단이나 발달 상태를 확진하기 위한 진단도구가 아니라, 국내 임상현장에서 발달지연이 의심되는 아동을 대상으로 조기 선별 및 전반적인 발달 상태를 확인하기 위한 목적으로 활용되는 선별검사라고 판단하였다.

동 검사의 활용 양상은 진료과 및 의료기관 유형, 진료 단계, 의료자원 여건, 대상 아동의 특성에 따라 달라질 수 있다고 보았다. 재활의학과에서는 발달지연 의심 아동의 전반적인 발달 상태를 확인하고 후속 평가 또는 개입 방향을 설정하는 과정에서 덴버발달검사를 활용하고 있는 것으로 파악되었다. 반면 일부 3차 의료기관의 소아청소년과 또는 정신건강의학과에서는 이미 국가건강검진이나 1차·2차 의료기관에서 선별이 이루어진 이후 내원하는 경우가 많고, 내원 초기부터 Bayley, WPPSI, WISC, CARS 등 정밀 진단·평가 도구가 우선적으로 적용되는 진료 흐름이 형성되어 있어 활용이 제한적일 수 있다고 보았다. 이에 따라 동 검사의 임상적 유용성은 단일한 기준이나 특정 발달 단계에 고정하여 판단하기보다, 실제 진료 흐름과 기관의 자원 여건을 고려하여 선택적으로 해석할 필요가 있다고 판단하였다.

또한 덴버발달검사는 국가건강검진(K-DST) 또는 다른 선별 과정에서 발달 위험 신호가 확인된 이후, 보호자 보고 기반 선별 결과를 임상가의 직접 관찰로 보완하거나 정밀평가 의뢰 여부를

판단하기 위한 중간 단계 평가 도구로 활용될 수 있다고 보았다. 아울러 정밀 지능검사 또는 발달검사를 즉시 시행하기 어렵거나 검사 소요 시간, 비용, 아동의 협조도 등 현실적 제약이 있는 경우, 아동의 현재 발달 상태를 확인하고 보호자 및 관련 기관에 설명하기 위한 참고자료로 활용될 수 있다고 판단하였다. 명확한 진단명이 없거나 경계선 수준의 발달 특성을 보이는 아동에서 치료 개입 필요성을 판단하거나, 발달재활서비스 등 중재 연계 과정에서 보조적 근거자료로 활용될 수 있다는 점도 함께 고려하였다.

아울러 동 검사의 임상적 위치와 활용 범위를 보다 명확히 평가하기 위해서는 실제 보건의료 현장에서의 사용 현황과 결과 활용 과정을 체계적으로 파악할 수 있는 자료의 축적·수집·분석이 필요하다고 판단하였다. 특히 발달재활서비스 의뢰 및 중재 과정과 연계하여 검사 실시 여부와 결과가 표준화된 형식으로 기록·관리될 수 있는 기반이 마련될 경우, 동 검사의 실제 임상적 역할과 유용성에 대한 보다 실증적인 평가가 가능할 것으로 보았다.

2026년 제2차 재평가전문위원회*(2026.2.6.)는 '덴버발달검사'에 대해 다음과 같이 심의하였다. 재평가전문위원회는 소위원회 검토 결과에 동의하였으며, 덴버발달검사가 스크리닝(screening) 또는 지속적 추적관찰(surveillance) 차원에서 발달 이상 여부를 확인하는 검사로서 갖는 의미와 지속 필요성에 대해서는 향후 추가적인 근거 검토를 통한 재평가가 필요하다는 의견을 제시하였다.

*「신의료기술평가에 관한 규칙」(보건복지부령 제1098호, 일부개정, 2025.9.7.시행) 개정으로 재평가전문위원회가 새로 구성되어 2025년 9월부터 운영되고 있다.

V 참고문헌

1. 국민건강보험공단. 비급여 정보 포털. Available from: https://www.nhis.or.kr/nbinfo/wbhfaa-06200m28.do?mode=view&articleNo=11000825.

1. 건강보험심사평가원 건강보험요양급여비용. 2024년 1월판. 서울; 2024.

2. 건강보험심사평가원 보건의료빅데이터개방시스템. Available from: https://opendata.hira.or.kr/home.do

3. 건강보험심사평가원 요양기관업무포털. Available from: https://biz.hira.or.kr/index.do?sso=ok

4. 대한소아신경학회. 발달지연. 소아신경질환 바로알기. Korean Child Neurology Society. 2023 Jul 20. Available from: https://www.cns.or.kr/bbs/diseases/list?category=C&utm

5. 대한소아재활·발달의학회. 소아재활의학. 3판. 서울: 군자출판사; 2021.

6. 보건복지부. 한국형 영유아 검진도구 독자 개발, 알기 쉬운 검진결과통보서 마련. 보도자료. 세종: 보건복지부; 2014 May 29.

7. 보건복지부. 건강검진 실시기준 [시행 2014.9.1.] [보건복지부고시 제2014-111호, 2014.7.22., 일부개정]. 세종: 보건복지부; 2014.

8. 일본 후생노동성. 진료보수점수표(2024년판). Tokyo: Ministry of Health, Labour and Welfare; 2024.

9. 질병관리본부. 한국 영유아 발달선별검사 사용 지침서. 청주: 질병관리본부; 2017.

10. Abessa TG, Worku BN, Kibebew MW, Valy J, Lemmens J, Thijs H, et al. Adaptation and standardization of a Western tool for assessing child development in non-Western low-income context. BMC Public Health. 2016;16:652.

11. American Academy of Pediatrics, Council on Children With Disabilities. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics. 2006;118(1):405-20.

12. American Medical Association. CPT Professional 2023 and E/M Companion 2023 Bundle. AMA Press; 2023.

13. Bayley N. Bayley scales of infant and toddler development. 3rd ed. Administration manual. San Antonio (TX): Harcourt Assessment; 2006.

14. Bayley N. Bayley Scales of Infant Development. 2nd ed. San Antonio (TX): The Psychological Corporation; 1993.

15. Chung HJ, Yang D, Kim GH, Kim SK, Kim SW, Kim YK, et al. Development of the Korean Developmental Screening Test for Infants and Children (K-DST). Clin Exp Pediatr. 2020;63(11):438-46.

16. Christovao IS, Pereira DAG, Magalhães LC, Camargos ACR. Predictive validity of the Denver developmental screening test (Denver-II) to detect risk of developmental coordination disorder in preterm children. Early Hum Dev. 2023;184:105836. doi:10.1016/j.earlhumdev.2023.105836.

17. De-Andrés-Beltrán B, Rodríguez-Fernández ÁL, Güeita-Rodríguez J, Lambeck J. Evaluation of the psychometric properties of the Spanish version of the Denver Developmental Screening Test II. Eur J Pediatr. 2015;174(3):325-9.

18. Del Rosario C, Slevin M, Molloy EJ, Quigley J, Nixon E. How to use the Bayley Scales of Infant and Toddler Development. Arch Dis Child Educ Pract Ed. 2021 Apr;106(2):108-112. doi: 10.1136/archdischild-2020-319063.

19. Frankenburg WK, Dodds JB. The Denver Developmental Screening Test. J Pediatr. 1967;71(2):181-91.

20. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics. 1992;89(1):91-7.

21. Ga H, Kwon JY, Kim Y, Lee J. Validity of the Korean Ages and Stages Questionnaires. J Korean Acad Child Health Nurs. 2011;17(2):79-88.

22. Glascoe FP, Byrne KE. The accuracy of three developmental screening tests. J Early Interv. 1993;17(4):368-79.

23. Jeon MC, Kim YH, Chung SY, Lee IG, Kim JW, Whang KT. Accuracy of Denver II in developmental delay screening. Ann Child Neurol. 1997;5(1):111-18.

24. Jeong SU, Kim GC, Jeong HJ, Kim DK, Hong YR, Kim HD, et al. The validity of the Bayley-III and DDST-II in preterm infants with neurodevelopmental impairment: a pilot study. Ann Rehabil Med. 2017;41(5):851-857. doi:10.5535/arm.2017.41.5.851.

25. Johnson S, Marlow N. Developmental screening and assessment tools in neurodevelopmental follow-up of preterm infants. Arch Dis Child Fetal Neonatal Ed. 2006;91(4):F289-93.

26. Lee K. Denver developmental screening II test and development of Seoul children. Korean J Pediatr. 1996;39(9):1210-4.

27. Rah SS, Jung M, Lee K, Kang H, Jang S, Park J, et al. Systematic review and meta-analysis: real-world accuracy of children's developmental screening tests. J Am Acad Child Adolesc Psychiatry. 2023;62(10):1095-1109.

28. Ranjitkar S, Kvestad I, Strand TA, Ulak M, Shrestha M, Chandyo RK, et al. Acceptability and reliability of the Bayley Scales of Infant and Toddler Development-III among children in Bhaktapur, Nepal. Front Psychol. 2018;9:1265. doi:10.3389/fpsyg.2018.01265.

29. Santos JAT, Ayupe KMA, Lima ALO, Albuquerque KA, Morgado FFDR, Gutierres Filho PJB. Propriedades psicométricas da versão brasileira do Denver II: teste de triagem do desenvolvimento. Cien Saude Colet. 2022;27(3):1097-106.

30. Shahshahani S, Vameghi R, Azari N, Sajedi F, Kazemnejad A. Validity and reliability determination of Denver Developmental Screening Test-II in 0-6 year-olds in Tehran. Iran J Pediatr. 2010;20(3):313-22.

31. Shin HS, Han KJ, Oh KS, Oh JJ, Ha MN. Denver developmental screening test: a cross-cultural comparison. J Korean Community Nurs. 2002;13(1):89-97.

32. Sondaar M, van Kessel BJM, de Kleine MJK, Briët JM, den Ouden AL, van Baar A. Do pediatricians recognize cognitive developmental problems in preterm children at age 5 years? J Dev Phys Disabil. 2008;20:21-9.

33. Srinithiwat B, Ularntinon S. Concurrent validity of the Ages & Stages Questionnaires, Third Edition, Thai-version (ASQ-3 Thai) with the Denver Developmental Screening Test II (DDST-II). J Med Assoc Thai. 2014;97 Suppl 6:S6-13.

34. Wijedasa D. Developmental screening in context: adaptation and standardization of the Denver Developmental Screening Test-II (DDST-II) for Sri Lankan children. Child Care Health Dev. 2012;38(6):889-99.

VI 부록

1. 의료기술재평가위원회

덴버발달검사 재평가를 위해 기존 의료기술재평가위원회(19명) 및 관련 법령 개정에 따라 새로 구성된 재평가전문위원회(20명)*가 총 2회 개최되었다.

*「신의료기술평가에 관한 규칙」(보건복지부령 제1098호, 일부개정, 2025.9.7. 시행) 개정으로 재평가전문위원회가 새로 구성되어 2025년 9월부터 운영되고 있다.

1.1 2025년 제7차 의료기술재평가위원회

- 회의일시: 2025년 7월 11일
- 회의내용: 재평가 프로토콜 및 소위원회 구성안 심의

1.2 2026년 제2차 재평가전문위원회

- 회의일시: 2026년 2월 6일
- 회의내용: 결론검토 및 최종심의

2. 소위원회

덴버발달검사 소위원회는 의료기술재평가 자문단에서 무작위로 선정된 이해상충 관계가 없는 임상전문가로 총 7인(소아청소년과 2인, 정신건강의학과(소아) 1인, 재활의학과(소아) 1인, 임상심리분석전문가 2인, 근거기반의학 1인)으로 구성하였다. 소위원회 활동 현황은 다음과 같다.

2.1 제1차 소위원회

- 회의일시: 2025년 9월 24일
- 회의내용: 평가계획서 논의

2.2 제2차 소위원회

- 회의일시: 2025년 11월 4일
- 회의내용: 연구선택 결과보고, 자료분석 계획 논의

2.3 제3차 소위원회

- 회의일시: 2025년 12월 10일
- 회의내용: 자료분석 결과보고, 결론방향 논의

2.4 제4차 소위원회

- 회의일시: 2025년 12월 22일~28일 서면 검토
- 회의내용: 결론 검토 및 제언 사항

3. 연구검색 현황

3.1 국외 데이터베이스

3.1.1 Ovid MEDLINE® 1946~현재까지

(검색일: 2025.10.14.)

구분	연번	검색어	검색결과(건)
P	1	(Toddler or Child or Preschool or Pre-school or Pre school or School-aged or School aged* or Elementary student or Baby or Babies or Kid or Kids).mp.	2,619,435
I	2	(Denver Development Screening Test or Denver Developmental Screening Test or Denver II Developmental Screening Test or Denver II Developmental Test).mp.	402
	3	(Infant & Child Monitoring Questionnaire or Korean Developmental Screening Test or Parent Questionnaire for the Developmental Evaluation or Parental Assessment Screening).mp.	32
	4	OR/2-3	434
P&I	5	AND/1-4	355

3.1.2 APA PsycInfo(via Ovid) 1806 to October 2025 Week 1

(검색일: 2025.10.14.)

구분	연번	검색어	검색결과(건)
P	1	(Toddler or Child or Preschool or Pre-school or Pre school or School-aged or School aged* or Elementary student or Baby or Babies or Kid or Kids).mp.	587,695
I	2	(Denver Development Screening Test or Denver Developmental Screening Test or Denver II Developmental Screening Test or Denver II Developmental Test).mp.	407
	3	(Infant & Child Monitoring Questionnaire or Korean Developmental Screening Test or Parent Questionnaire for the Developmental Evaluation or Parental Assessment Screening).mp.	6
	4	OR/2-3	413
P&I	5	AND/1-4	284

3.2 국내 데이터베이스

(검색일: 2025.10.14.)

데이터베이스	연번	검색어	검색결과(건)
KoreaMed	1	Denver	118
	2	DDST	19
	3	Child Developmental Screening	114
	소계		251

4. 최종선택연구

연번	서지정보
1	Christovao IS, Pereira DAG, Magalhães LC, Camargos ACR. Predictive validity of the Denver developmental screening test (Denver-II) to detect risk of developmental coordination disorder in preterm children. Early Hum Dev. 2023;184:105836. doi:10.1016/j.earlhumdev.2023.105836.
2	Santos JAT, Ayupe KMA, Lima ALO, Albuquerque KA, Morgado FFDR, Gutierres Filho PJB. Psychometric properties of the Brazilian version of the developmental screening test: Denver II. Cien Saude Colet. 2022;27(3):1097-1106.
3	Sondoar M, van Kessel B, An K, Kleine MJK, Pires J, Ouden AL, et al. Do paediatricians recognize cognitive developmental problems in preterm children at age 5 years? J Dev Phys Disabil. 2019;31:311-321.
4	Jeong SU, Kim GC, Jeong HJ, Kim DK, Hong YR, Kim HD, et al. The validity of the Bayley-III and DDST-II in preterm infants with neurodevelopmental impairment: a pilot study. Ann Rehabil Med. 2017;41(5):851-857. doi:10.5535/arm.2017.41.5.851.
5	Abessa TG, Worku BN, Kibebew MW, Valy J, Lemmens J, Thijs H, et al. Adaptation and standardization of a Western tool for assessing child development in non-Western low-income context. BMC Public Health. 2016;16:652.
6	De-Andrés-Beltrán B, Rodríguez-Fernández AL, Güeita-Rodríguez J, Lambeck J. Evaluation of the psychometric properties of the Spanish version of the Denver Developmental Screening Test II. Eur J Pediatr. 2015;174(3):325-329.
7	Wijedasa D. Developmental screening in context: adaptation and standardization of the Denver Developmental Screening Test-II (DDST-II) for Sri Lankan children. Child Care Health Dev. 2012;38(6):889-899.
8	Ga H, Kwon JY, Kim Y, Lee J. Validity of the Korean Ages and Stages Questionnaires. J Korean Acad Child Health Nurs. 2011;17(2):79-88.
9	Shahshahani S, Vameghi R, Azari N, Sajedi F, Kazemnejad A. Validity and reliability determination of Denver Developmental Screening Test-II in 0-6 year-olds in Tehran. Iran J Pediatr. 2010;20(3):313-322.
10	Srinithiwat B, Ularntinon S. Concurrent validity of the Ages & Stages Questionnaires, Third Edition, Thai-Version (ASQ-3 Thai) with the Denver Developmental Screening Test II (DDST-II) in developmental screening of 18, 24, and 30 months old children at Queen Sirikit National Institute of Child Health. J Med Assoc Thai. 2014;97 Suppl 6:S6-S13.
11	Glascoe FP, Byrne KE. The accuracy of three developmental screening tests. J Early Interv. 1993;17(4):368-379.
12	Shin HS, Han KJ, Oh KS, Oh JJ, Ha MN. Denver developmental screening test: a cross-cultural comparison. J Korean Community Nurs. 2002;13(1):89-97.
13	Jeon MC, Kim YH, Chung SY, Lee IG, Kim JW, Whang KT. Accuracy of Denver II in developmental delay screening. Ann Child Neurol. 1997;5(1):111-118.
14	Lee K. Denver developmental screening II test and development of Seoul children. Korean J Pediatr. 1996;39(9):1210-1214.
15	Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics. 1992;89(1):91-97.

NECA 한국보건의료연구원
National Evidence-based Healthcare Collaborating Agency

발행일 2026. 6. 30.
발행인 이 재 태
발행처 한국보건의료연구원

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

ISBN : 979-11-7337-198-1