

# 형사미성년자 연령을 둘러싼 증거기반(evidence-based) 연구의 검토와 전망\*

박성훈\*\* · 박신의\*\*\*

## 국 | 문 | 요 | 약

이 논문은 최근 정치사회적으로 큰 이슈가 되고 있는 소년범죄자의 형사처벌에 관심을 가지고, 국내 선행연구 및 언론 내용을 분석하여 주요 쟁점을 정리하였다. 그러나 국내 선행연구의 결과가 규범적 주장에 그치거나 근거로 제시된 경험적 증거가 오히려 논란이 되고 있다는 사실에 주목하여 최근 널리 받아들여지고 있는 '증거기반' 개념을 재검토하고 증거기반 형사정책의 역사와 의미를 살펴보았다. 특히, 미국의 연구 사례를 바탕으로 소년의 책임능력 인정에 관한 뇌과학 분야의 증거, 그리고 소년사범의 형사이송 효과에 관한 증거를 체계적으로 검토하였다. 결론적으로 이 논문은 미국의 연구 사례가 현재 국내에서 벌어지고 있는 형사미성년자 연령 논쟁과 관련하여 시사점을 줄 수 있으나, 직접적 증거가 되기 어려울 뿐 아니라 증거의 효과 역시 분명하지 않다는 점, 증거기반 연구에 내재한 탈맥락화와 방법론적 획일화 등 증거기반의 한계를 인식하는 것이 중요하다는 점을 지적하였다. 다만, 이러한 한계에도 불구하고 국내 형사정책 분야에서 증거기반 연구를 통한 타당한 지식의 축적과 이에 기반한 정책 수립은 계속해서 지향해야 할 것으로 전망하였다.

DOI : <https://doi.org/10.36889/KCR.2023.9.30.3.1>.

❖ 주제어 : 소년범죄, 촉법소년, 형사미성년자, 증거기반정책, 랜덤통제실험

\* 이 논문은 2023년 6월 9일 개최한 한국형사·법무정책연구원과 대한범죄학회 공동학술대회 발표문을 수정·보완한 것으로, 유용한 논평을 해 주신 익명의 심사위원들께 감사의 말씀을 전한다.

\*\* 한국형사·법무정책연구원 선임연구위원, 사회학박사

\*\*\* 신시내티대학교 형사사법학과 박사과정

## I. 서론

2022년 10월 26일 법무부는 「소년법 일부개정법률안」 내용 중 ‘촉법소년 연령 기준 현실화’를 제안하였다. 촉법소년 연령 상한을 13세 미만으로 하향하고, 형사처벌이 가능한 소년의 연령을 13세로 낮추자는 주장이다. 그 근거로 최근 촉법소년 범죄 증가 및 수법의 흉포화, 촉법소년 제도의 악용 사례 등으로 인해 촉법소년 연령 상한을 현실화해야 한다는 여론이 우세하다는 점, 흉포화된 소년범죄로부터 국민을 보호할 필요가 있다는 점, 보호처분을 받은 촉법소년 중 13세 소년의 비중이 증가하는 점을 들고 있다. 그러나 처벌 위주의 일방적 조치만으로 소년범죄자의 재범을 막을 수 없으므로 맞춤형 보호 처분 마련, 소년보호절차 개선, 피해자 보호 강화, 형사사법 절차에서 전문성 제고, 소년 범죄통계 관리시스템 구축 등 소년사법체계를 전반적으로 정비하는 것을 제안하였다(법무부, 2022). 법무부의 제안에 대해 대한변호사협회는 “촉법소년의 연령 조정은 타당하다”는 의견을 제시하였다. 그 근거로는 2007년 「소년법」 개정논의 당시에도 촉법소년 상한 연령인하에 대한 주장이 있었으나 받아들여지지 않았고, 개정 이후 15년 동안 급속하게 바뀐 사회변화를 고려할 필요가 있다고 지적하였다. 다른 근거로는 촉법소년 기준 연령 인하가 보호처분이 꼭 필요한 소년에 대한 관심과 지원을 집중시켜 범죄원인과 환경에 대한 다각적인 처우를 기반으로 하는 「소년법」 취지를 살릴 수 있다고 하였다(대한변호사협회, 2022).

그러나 법원행정처는 촉법소년 기준 연령 개정안에 반대 의견을 표명하였다. 그 근거로 13세 소년이 ‘형사책임능력’을 갖췄다고 단정 짓기 어렵다는 점, 현행법상 13세 소년에 대해 부과되는 보호처분이 형사처벌과 비교하여 결코 경미하다고 볼 수 없는 점, 13세 소년이 형벌 법령에 저촉되는 행위를 하게 된 원인으로 지적되는 소년의 가정환경 개선이나 정신질환 치료 등 적극적인 사회적 지원 없이 촉법소년 연령을 낮추는 것만으로 근본적인 해결이 어렵다는 점을 이유로 들었다(법률신문, 2023/02/23). 촉법소년 연령에 대한 법원의 입장은 국가인권위원회가 제시한 ‘형사미성년자 연령하한 및 촉법소년 상한 연령하향 조정’에 대한 결정문에 상세히 기술되어 있다. 국가위원회는 대한민국헌법과 국제인권기준을 근거로 형사미성년자 연령을 14세에서 12세 또는 13세로 낮추려는 의원 법률안과 법무부 입장에 대해 소년범죄 예방과 재범 방지를 위한 실효적 대안으로

바람직하지 않다는 의견을 표명하였다. 소년범죄자는 성인보다 범죄의 상습성 정도가 약하여 개선가능성이 크다는 점, 소년보호 이념에 근거하여 전 세계적으로 소년의 비행이나 범죄에 대해서는 일반적인 형사절차와 병행하거나 독립된 처리 절차를 두고 있다는 점이 근거로 제시되었다. 국가인권위원회는 아동의 존엄과 권리를 보장하는 유엔 아동권리협약 제40조를 근거로 유엔 아동권리위원회가 대한민국 제5·6차 국가보고서(CRC/C/KOR/CO/5-6)에서 “형사책임 최저연령을 14세로 유지하고, 14세 아동을 범죄자로 취급하거나 구금하지 않을 것을 권고하였다”는 사실도 언급하였다(국가인권위원회, 2022).

2007년 「소년법」의 개정을 통해 한 차례 소년범죄에 대한 처벌을 강화하였지만, 여전히 형사책임연령에 대한 여론은 차갑고, 엄벌을 지향하는 목소리는 유지되고 있다. 한국리서치가 전국의 성인 남녀 1,000명을 대상으로 조사한 결과, ‘촉법소년의 범죄가 심각하다’는 의견이 95%였고, 촉법소년 연령 하향에 대해서는 ‘지금보다 더 낮춰야 한다’는 의견이 96%로 나타났다.<sup>1)</sup> 사실 형사미성년자 연령을 조정하지는 주장은 비단 한국뿐 아니라 전 세계에서 꾸준히 제기되는 논의 중 하나이다. 소년범죄자 처벌에 대한 사회적 요구는 갈수록 높아지고 있으나, 관련법률을 개정하거나 여론의 요구를 쉽게 반영하지 못하는 이유는 소년법 연령하향 문제가 형사책임주의 목적과 소년보호 이념이 교차하는 지점에 있기 때문이다(이덕인, 2012). 엄벌만이 능사는 아니라는 점은 서구의 많은 연구에서 자명하게 드러나고 있기 때문에(Cullen and Jonson, 2016) 새로 도입하려는 법안 및 제도의 효과, 그 실익에 대한 평가와 검증이 필요하다는 주장이 제기될 수밖에 없다.

국내 연구자들은 형사미성년자 연령하향 문제를 다루면서 규범적 논거, 공식통계에 대한 해석, 사회발전에 따른 연령의 의미변화, 이론적인 논의 등을 검토하여 찬성 혹은 반대의 주장을 펼쳐왔다. 그러나 국내에서 관련 주제에 관한 실증적이고 경험적 연구가 충분히 수반되지 못하면서, 논의가 더 나아가지 못하고 제자리를 맴도는 경향이 없지 않다. 이러한 답답한 상황을 극복하기 위해 널리 알려진 ‘증거기반’ 연구 및 정책의 필요성이 다시금 강조되고 있다(박철현 2014; 하태훈, 2022). 이 논문은 국내에서도 증거기반 관점을 가지고 형사미성년자 문제를 바라본다면 정체된 논의가 조금 더 진전된 방향으로 진행될 수 있을 것이라는 기대에서 비롯되었다.

1) 한국리서치 주간 리포트(제215-2호) 여론 속의 여론 “촉법소년 범죄 해결 여론”(2023년 1월 25일).

이 논문은 먼저 소년범죄 연령 하향을 둘러싼 찬반 양측의 논점을 정리하고, 국내 소년사법정책 분야에서 필요하다고 여겨지는 증거기반의 개념과 의미를 재검토해 보고자 한다. 이어서 소년사법정책과 관련하여 증거기반 관점에서 수행된 미국의 연구사례를 살펴보고자 한다. 미국의 연구사례는 소년의 책임연령(doli capax)에 대한 과학적 증거를 다룬 연구와 형사이송제도(criminal transfer)의 효과성을 증거기반 관점에서 다룬 연구로 나누어 상세히 살펴볼 것이다. 미국의 사례에서 제시된 증거를 통해 향후 국내에서 증거기반 연구를 수행하기 위해 필요한 것이 무엇이고, 증거기반이 가지고 있는 비판과 한계도 살펴봄으로써 소년사법정책을 비롯한 국내 형사정책 분야에서 증거기반의 가능성을 전망해 보고자 한다.

## II. 형사미성년자 연령을 둘러싼 논쟁

「소년법」 상의 촉법소년 연령 및 「형법」 상 형사미성년자 연령 하향을 둘러싸고 법무부, 법원, 대한변호사협회, 국가인권위원회 등 관련 기관 간 논의가 팽팽하게 진행되는 가운데, 학계의 동향은 대체로 연령기준 하향에 반대하는 논의가 대부분이다. 다만, 최근 학계에서도 소년범죄자에 대한 처벌 강화와 함께 연령기준 하향에 대한 필요성 혹은 재검토에 대한 논의가 조금씩 증가하고 있다(손현중·성진기, 2023: 127). 연령기준 하향을 주장하는 학자들은 변화된 사회환경, 공공의 안전확보, 특별예방효과 등을 근거로 제시한다. 형사미성년자 연령 및 촉법소년 연령의 하향을 둘러싼 국내 학계의 쟁점은 크게 소년범죄통계 추세에 대한 해석, 소년범죄자의 책임능력에 대한 해석으로 대별할 수 있다.

먼저, 소년범죄통계 추세와 관련하여 연령 하향을 주장하는 측에서는 소년범죄가 갈수록 증가하고 있고, 강력범죄를 저지르는 연령 또한 낮아지는 추세이며, 범죄의 양상도 성인 못지 않게 잔혹해지고 있다는 점을 지적한다(점승현, 2022). 이러한 주장은 우리나라 경제수준의 향상, 인터넷 등 각종 미디어의 보편화에 따라 소년들의 신체적, 정신적 성장이 과거에 비해 더 빨라졌을 뿐 아니라 어릴 때부터 법에 대한 해박한 지식을 갖추고 있어 과거와 같은 처분만으로는 시대적 변화를 반영하기 어렵다는 주장으로 이어진다

(박지혜·이수정, 2022). 2015년 10월 갯마을사건을 비롯하여 2017년 인천초등학생 살인 사건, 부산여중생폭행사건, 2021년 양산여중생폭행사건 등 국민적 공분을 불러일으킬 만한 소년범죄사건이 연이어 발생하면서 언론과 여론은 소년범죄의 흉포화에 상응하는 법령 개정의 필요성에 더 공감하고 있다. 정치권 역시 경찰, 검찰, 법무부, 법원 등 형사 사법기관이 제공하는 통계를 근거로 촉법소년이 저지르는 강력범죄 사건이 증가했다고 주장하면서 연령 하향의 입법 필요성을 주장한다(박찬걸, 2020). 일례로 국민의 힘 소속의 의원은 경찰이 제시한 촉법소년 범죄집수 현황 자료를 근거로 최근 5년간 촉법소년 강력범죄 소년부 송치 건수는 2017년 6,286건에서 2021년 8,474건으로 35%가량 증가했다면서, 촉법소년의 범죄가 저연령화·흉포화되는 만큼 현실적인 연령기준의 필요성을 주장하였다(서울경제, 2022/04/01). 검찰이 발표하는 『범죄분석』을 근거로 “2006년 소년사건 재범율이 28.9%에서 2016년 42.6%로 증가했다”(박상식, 2017)는 주장, 그리고 경찰·검찰·법원에서 발표한 자료를 비교·분석한 결과 “14세 미만의 범죄가 증가하고 있음을 확인할 수 있다”는 주장은 여러 해석의 차이에도 불구하고 재범을 포함한 “소년범죄가 줄고 있다고 보기 어렵다”(박지혜·이수정, 2022)는 근거로 제시된다.

반면, 동일한 자료를 검토한 결과 소년범죄의 전반적 추세는 감소하고 있고 증가세 역시 16~17세에서 두드러진 것이므로 굳이 촉법소년의 연령과 형사미성년자의 연령을 낮추자는 것은 설득력이 없다는 반론도 제기된다(박찬걸, 2013). 2016년 14세~15세의 흉악범죄가 32.0%로 전년 21.8%에 비하여 증가한 것은 사실이나, 2009년(30.9%) 이후 10년간 변화 추세를 보건대 증감을 반복하는 범위를 벗어나지 않는다는 점(이용식, 2017; 박수희·문준섭, 2018), 2020년 창궐했던 코로나 팬데믹이 촉법소년의 범죄에 미친 영향은 확인하지 않은 채 코로나 이전과 이후의 통계를 단순히 비교하는 것은 적절하지 않다는 점(원혜옥, 2022), 현재 국내에서 공표되는 통계는 촉법소년 사건 발생 현황을 전체적으로 명확하게 보여주지 못하므로 저연령화·흉포화 주장이 객관적으로 타당한지를 판단하기가 어렵다는 점(국가인권위원회, 2022)이 반론의 근거로 제시되고 있다.

다음으로 형사미성년자의 책임능력에 관한 논쟁을 살펴보자. 연령하향을 주장하는 입장에서는 요즘 청소년들이 육체적 성장뿐 아니라 정신적으로도 성숙하여 1953년 소년법 제정 당시와 다르기 때문에 변화한 사회적 흐름에 맞추어 기준을 바꾸는 것이 합리적이라고 본다. 오히려 14세 미만 형사미성년자를 일률적으로 처벌하지 않는 것이야말로 불

합리하다는 것이다(정신교·차시환, 2014; 김두상·박상식, 2016). 이들은 최근 소년범죄와 성인범죄 간 질적 차이가 점차 없어지는 것을 고려할 때, 형사책임능력을 연령에 따라 획일적으로 적용할 것이 아니라 최소한 10세이상 19세미만 소년범에 대해서는 범죄 소년이나 촉법소년 구분을 폐지하고, 범죄의 성격과 질에 따라 보안처분 및 형벌의 부과가 가능하도록 해야 한다(김봉수, 2013: 191)고 주장한다. 「형법」 제9조(형사미성년자)의 규정이 연령 기준을 획일적으로 설정한 것은 명확성과 법적 안정성을 기하기 위한 것으로 형사책임연령을 적극적으로 규정하는 조문은 아니라는 해석도 있다. 14세에 도달하지 않더라도 실질적 의미에서 사물변별능력이나 의사결정능력을 갖춘 경우가 많고 성장했다고 하더라도 책임능력이 갑자기 생기는 것은 아니기 때문에 「형법」 제9조는 일정 연령에 도달하지 않은 자의 책임능력 없음을 의제함으로써 그들에게 형벌부과가 불가능하다는 점을 소극적으로 규정하고 있을 뿐이라는 것이다(남선모, 2019: 36).

그러나 소년범의 책임능력 인정에 대한 반론도 만만치 않다. 1953년 이래 비록 소년의 신체적 성장과 정신적 성숙이 빨라졌다고 하더라도 14세 미만자에게 일률적으로 형사책임을 지을 만큼 그들이 충분히 성숙했는지는 판단할 수 없다는 것이다(박찬걸, 2020: 155). 형사책임능력과 관련하여 기준연령의 변천사와 비교법적 검토를 다룬 연구에서는 과거와 다르게 점차 빨라지는 소년들의 정신적·육체적 성장 속도와 저연령 소년범죄의 심각성으로 인해 책임연령을 낮추어야 한다는 여론이 지배적이지만 촉법 연령층이 최근 갑자기 증가한 것은 아니며, 이 같은 인식의 근거에는 형사법을 지탱하는 책임주의를 흔들고 성인이든 소년이든 범죄자를 사회악으로 간주하는 ‘적대형법(Feindstrafrecht)’ 논리가 숨어 있다는 점을 지적한다(이덕인, 2012: 27). 다른 연구에서는 형사미성년 제도가 연령을 기준으로 형벌로 책임을 물을 수 있는 시기와 형벌에 의한 책임비난이 타당하지 않은 형사미성년의 한계를 규정한 것이므로, 형법의 목적과 논리상 형벌로 예방효과를 기대할 수 없는 사람에게 형벌을 통해 비난하는 것이 허용되지 않으며, 형벌 이외의 제재(소년법의 보호처분)를 통해 더 유효한 특별예방효과를 기대하는 것이 타당하다고 주장한다(김연지, 2018: 175). 형사책임능력은 형법상 책임주의와 소년법상 보호주의가 교차하는 영역에 자리하고 있기 때문에 연령에 따라 일률적으로 책임능력을 판단하는 것은 충분히 타당한 근거를 가지고 논의해야 한다는 것이다(김연지, 2018). 이러한 주장의 선상에서 최근 청소년 뇌과학 연구를 기초로 소년의 ‘미성숙

한 뇌’는 엄연한 ‘사실’이며, 미국 연방대법원과 법원이 청소년의 뇌과학분야 연구를 상당히 수용하여 소년범의 변별능력과 의사결정능력이 제대로 작동하지 못한다면 그 감정을 인정할 수 있다는 논리도 반대 근거로 제시되었다(이인영, 2015).

지금까지 살펴본 논쟁을 정리하면, 대체로 연령하향에 찬성하는 측은 소년범죄통계에 기초하여 저연령화·홍포화에 대응하고, 소년 재범 증가를 막기 위해서라도 형사처분이 가능한 연령을 낮추어야 한다고 주장한다. 반면, 연령하향에 반대하는 측에서는 성인과 다른 소년 특수성을 인정한 소년보호이념 및 국제인권기준, 그리고 뇌과학에 대한 최근 연구성구를 바탕으로 소년범의 형사책임능력을 인정하기 쉽지 않다는 반론을 개진하고 있다.

이와 다르게 연령하향에 반대하면서 형사책임능력보다는 미국에서 이미 시행된 엄벌주의 소년사법정책의 부정적 효과를 소개한 연구도 있다(박미량 2014; 박선영, 2019). 두 연구 모두 저자가 범죄학자라는 점, 1980년대 우리나라와 유사한 논쟁을 거쳐 엄벌주의 소년사법제도를 시행한 미국의 사례를 다루고 있다는 점, 형사이송 혹은 형사처벌이 소년범죄의 억제에 미치는 경험적 연구를 소개했다는 점에서 공통적이다. 박미량(2014)의 연구가 미국의 다양한 소년범 형사이송제도를 상세하게 설명하고 있다면, 박선영(2019)의 연구는 정책의 평가 결과를 일반억제, 특수억제, 재범율 등 여러 측면에서 살펴보고 있다는 점에서 차이가 있다. 두 연구는 소년사범에 대한 국내 논의가 다소 불분명한 통계자료나 규범적인 정당성 논란에 몰입되어 있다는 점에 착안하여 법률의 개정 및 제도의 도입이 실제로 범죄현상에 미치는 결과를 살펴보았다는 점에서 주목할만하다. 우리보다 앞서 도입된 엄벌주의 소년사법 정책의 검토 결과는 “(정책의) 효과성을 추적할 수 있는 데이터 시스템을 의무적으로 구축”하자는 결론(박미량, 2014: 125), 그리고 “엄벌주의가 과연 과학적이고 객관적인 증거자료에 기반한 주장인지, 실제로 범죄억제를 이끌 수 있는 효과적인 대응 방안인지 냉철하고 심도깊은 분석과 논의를 해야 할 필요”가 있다는 결론(박선영, 2019: 106)으로 마무리되고 있다.

사실 이러한 주장은 앞서 소개한 국내 여러 학자도 언급한 바 있다. 이들의 공통 의견은 언론에 보도되는 잔혹한 사건으로 대중의 불안과 분노가 높아지고, 정치권은 이를 신속히 입법화하여 대응하는 정책 결정은 매우 비과학적이므로 정확한 실태조사와 실증연구가 필요하며, 새로운 정책과 법률은 그것이 미칠 영향에 대한 평가작업을 거친 후 결

정해야 한다는 것이다. 무엇보다 뜨거운 논쟁 가운데 위치한 “소년사법정책이야말로 증거기반 형사정책이 적극적으로 도입되어야 할 분야”(하태훈, 2022: 28)라는 주장이 공감을 얻고 있다. 소년사범에 대한 형사처벌의 효과성을 진단할 수 있는 국가 공식통계나 평가연구가 부재한 상황에서 소년법의 폐지를 논의하고, 형사책임연령을 인하하는 것은 증거에 기반한 정책이 아닌 감정적 대응이 될 수 있기 때문이다(박선영, 2019: 121).

### Ⅲ. 증거기반의 역사와 의미

#### 1. 증거기반의 개념과 특성

널리 알려진 대로 ‘증거기반’에 대한 초기 관심은 보건의학 분야에서 시작하였다. 1972년 스코트랜드 의사 코크란(Archie Cochrane)은 당시 의사들이 개인적 임상경험에 근거하여 처방하는 관행을 비판하면서 객관적이고 타당한 근거에 기초한 의료행위 필요성을 주장했다(박철현, 2014). ‘증거기반의학(evidence-based medicine)’이라는 용어가 문헌에 처음 등장한 시점은 1991년 캐나다 McMaster 대학 Gordon Guyatt의 논문에서다. 그는 빈혈이 의심되는 환자에게 쓰이는 진단 방법을 예로 들었다. Guyatt는 기존의 무차별적 검사에 의한 진단과정과 민감도·특이도 등 정량적 데이터를 고려한 진단과정을 비교하면서 후자와 같은 과학적 근거에 기반을 둔 진료가 필요하다고 주장했다. 이후 증거기반의학은 Guyatt 등이 연구모임(evidence-based medicine group)을 결성하면서 크게 발전하였다(김수영, 2001).

증거기반의학 발전에 이바지한 또 다른 학자인 David Sackett(1995)은 증거기반의학을 다음과 같이 정의하면서 그 필요성을 강조하였다.

개별 환자에게 서비스를 제공하든 집단에 서비스를 제공하든 우리 의사들은 항상 가능한 최선의 근거에 기반하여 결정하고 행동하려고 노력해 왔다. 무작위 임상시험(randomized trial) 등장은 진단, 예후 및 치료에 대한 임상적 근거를 확립하는 방식에 근본적 변화를 예고하고 있다. 이처럼 빠르게 성장하는 증거들을 따져보고 타당성과 유용성 측면에서 비판적으로 평가하여 실제 임상에 적용하는 것을 ‘증거기반의학(evidence-

based medicine)’이라고 부른다(Sackett & Rosenberg, 1995: 620).

그에 따르면, 증거기반의학(EBM)이 등장하게 된 첫 번째 배경은 최신 의학정보가 너무 많다는 점이다. 대부분 교과서는 발간과 더불어 과거의 지식이 되는데, 새롭게 쏟아지는 엄청난 양의 논문을 모두 이해하고 따라가는 것은 불가능하다. 이런 상황에서 의사는 임상적으로 문제가 발생하면 문헌을 찾기보다 주변 의사에게 묻는 경우가 많다. 하지만, 임상정보는 전문가마다 혹은 기관마다 다르게 평가할 수 있다. 둘째로 의사의 진료행위 가운데 상당수는 뒷받침할만한 근거가 뚜렷하지 않다는 점이다. Millenson(1997)은 대략 85% 의료행위가 객관적 근거가 없으며, 대부분 의사는 2,500개나 되는 의학저널을 거의 읽지 않고, 상당수 연구가 실험설계도 거치지 않은 결과라는 점을 지적하였다. 이를 보완하기 위해 의사 연수교육이 대안으로 제시되었으나 연수교육의 효과성 분석 결과는 의사의 지식향상에 도움을 주지 못한 것으로 나타났다(Sackett and Rosenberg, 1995; 김수영, 2001; Sherman, 2009). 의료·보건 분야의 증거기반은 과학적으로 효과성이 입증된 치료를 받아야 한다는 환자의 권리인식 증대, 효과가 불확실한 일부 의료기술에 대한 우려, 의료·보건 자원의 합리적인 사용에 대한 필요성, 전문가의 의견에 전적으로 의존하는 의사결정의 한계 등 여러 문제가 복합적으로 대두되면서 크게 공감을 얻을 수 있었다(서홍란, 2014).

증거기반의학은 개별 환자의 문제에 관해 결정을 내릴 때 세심하고 주의 깊게 최신 의학지식을 적용하는 것을 강조하며(Sackett et al., 1996: 71), 개인의 임상경험과 체계화된 연구에서 얻은 임상적 근거 중에서 최선의 것을 통합하여 환자 개인에게 적용하는 것을 목표로 한다. 여기서 ‘최신’은 의학지식이 발달하면서 기존 의료정보나 치료방법은 쇠퇴하기 때문에 이에 대처할 수 있어야 한다는 것이다. ‘개인의 임상경험’은 의사 개인이 임상경험을 통해 획득한 기술과 판단능력으로 진단, 환자의 상태, 권리, 선호에 대해 고려하고 결정하는 능력을 말한다. ‘최선의 근거’란 진단, 예후, 치료, 예방의 효과가 과학적으로 입증된 것을 사용해야 한다는 것을 말하고, ‘환자 개인’은 각 환자의 문제 해결에 필요한 정보를 수집하여 평가하고 실제 의료에 적용할 수 있어야 한다는 것을 의미한다. 증거기반의학에서는 ‘결과(outcome)’라는 용어에 대한 이해가 필요하다. 전통적으로 의사들은 결과를 생리학적 현상으로 생각하였다. 일례로 허혈성 심질환이 있는 환자에게 항부정맥 치료가 적절한지는 부정맥이 얼마나 억제되는지 여부로, 전립선비대증 환자가

약물치료를 받을 시 적절한 목표는 전립선 크기가 감소되는 것으로 이해했다. 이런 생리적, 해부학적 결과를 ‘대리결과(surrogate outcome)’라 하는데, 이런 것들이 증상해소, 일상기능회복, 생존과 같이 환자가 가장 많이 관심을 가지는 ‘최종 결과(end outcome)’와 밀접한 관련이 있을 것이라 가정한다. 그러나 실제로는 대리결과가 최종 결과를 대신 할 수 없는 경우가 많다. 예를 들어, 심근경색 후 심실부정맥 예방에 적절한 약을 처방하면 부정맥을 90%까지 억제할 수 있으나 위약에 비해 사망률은 2배 증가한다. 양성 전립선 비대증에 적절한 약을 투여하면 전립선의 크기는 감소하지만 빈뇨, 질박뇨의 증상개선에는 도움이 되지 않는다. 이러한 예들은 대리결과로 치료의 효과를 판단할 때 발생할 위험성을 보여준다. 치료의 목표가 통증 해소, 일상기능 회복, 생존 연장이라면 생리적 변화를 통한 유추보다 최종결과를 직접 측정해야 한다. 그러므로 증거기반의학이 말하는 결과란 대리결과가 아닌 최종결과를 의미한다(김수영, 2001: 3-4).

정리해보면, 증거기반의학이 말하는 증거기반의 특성은 해당 분야에서 타당성이 인정된 최선의 지식, 랜덤통제실험(randomized controlled trial, RCT)과 같이 엄격한 방법론에 기초한 최선의 증거, 의사(공급자)보다 환자(수요자) 중심의 가치 평가 등으로 요약할 수 있다.

## 2. 증거기반과 형사정책

19세기 이래 20세기 중반까지 ‘사회에 대한 과학(science of society)’으로 인간 행위에 관한 일반법칙을 제시할 것처럼 보였던 사회과학은 21세기로 접어들면서 점차 사람들의 기대에 부응하지 못하는 것으로 여겨졌다.<sup>2)</sup> 특히 정책 입안자들은 사회과학이 복잡한 현실을 깔끔하게 설명하기가 어려울 뿐 아니라 정책결정자와 시민들에게 적절한 해답을 제공할 수 없다고 생각하게 되었다.<sup>3)</sup> Stoker와 Evans(2016)에 따르면, 이런 상황에

2) 2007/08년 금융위기 이후 런던 정경대학(the London School of Economics)을 방문한 영국 여왕은 왜 경제학자들이 임박한 재정 문제를 예측하지 못했는지 따져 물었다(Pierce, 2008: Stoker & Evans, 2016에서 재인용).

3) 1950년대와 1960년대 미국에서 사회과학적 지식을 바탕으로 하여 인류 사회를 개선하고자 하는 정책들이 적극적으로 시행되기도 하였다. 하지만, 사회과학 지식에 바탕을 둔 정책을 통해 사회를 개선할 수 있다는 낙관적인 기대는 ‘Nothing Works(Martinson, 1974)’로 대변되는 회의적인 연구결과가 1970년대에 들어 잇따라 발표되면서 크게 꺾였다(한승훈·안혜선, 2021: 291).

직면한 사회과학이 제시한 방안은 두 가지이다. 첫째는 ‘모범사례 연구(best practice research)’로 새로운 연구를 제시하기보다 기존 연구를 선별하여 전달하는 방식이다. 대표적인 예가 특정 주제에 대한 유용한 증거를 모아 평가하는 체계적 검토(systematic reviews)이다. 둘째는 사회과학 연구를 정책과 연결하는 ‘증거기반 정책결정(evidence-based policymaking)’을 지원하는 것이다.

사회과학이 제시한 정책이 의도와 다르게 매번 실패한다는 비판은 상관성(correlation)에 기초한 낮은 과학적 기준에 초점이 맞추어졌고, 그 결과 인과성(causality)에 기초한 양적 방법과 실험설계 등 사회과학의 ‘과학성’을 높여야 한다는 주장으로 전개되었다. 1990년대 후반부터 영국, 호주, 미국 등 영미권 국가를 중심으로 일어난 ‘증거기반정책’에 대한 관심은 재정적자 타개 및 지지율 확보를 위한 정치권의 이해와 맞물리면서 폭발적으로 증가하였다. 증거기반에 관한 사회과학적 논의는 공중보건학, 교육학, 심리학, 사회복지학, 범죄학 등 사람들을 대상으로 한 정책 분야로 확대되면서 큰 기대감을 불러일으켰다(임민경 외, 2013; 서홍란, 2014; 오철호, 2015; 김병배, 2017; 한승훈·안혜선, 2021).

증거기반은 형사정책 분야에서 증거기반 치안정책(evidence-based policing, Sherman 1998), 증거기반 교정정책(evidence-based corrections, MacKenzie 2000), 증거기반 범죄예방(evidence-based crime prevention, Welsh and Farrington, 2001) 등으로 다양하게 확대되었다. 형사정책에 증거기반을 도입하고 후일 실험범죄학(experimental criminology)으로까지 발전시킨 Sherman(2009)은 의학과 마찬가지로 형사정책도 랜덤통제실험과 인과성에 기초한 증거를 축적할 필요가 있다고 주장하였다. 애초 Sherman(1998)이 증거기반의 중요성을 언급한 분야는 경찰활동, 즉 치안정책 분야였다. 그는 증거기반 의학을 차용하여 증거기반 치안정책을 “경찰 업무에 관한 최상의 연구를 활용하여 (정책의 목적을) 이행할 수 있는 지침을 만들고 그 결과를 평가하는 것”으로 정의하고, 가정폭력에 대한 경찰 대응 연구를 예시로 들었다. 그는 증거기반 치안정책을 통제된 조건에서 의도한 정책이 제대로 실행되었을 때 최선의 방안을 도출하는 ‘기초 연구(basic research)’, 기초 연구에서 발견한 방안을 실제 현장에 적용하여 각 조직이 달성한 결과를 지속해서 분석하는 ‘성과 연구(outcomes research)’로 구분하였다(Sherman, 1998: 3-4).

교정 분야에서도 유사한 주장이 제기되었다. Sherman의 영향을 받은 Mackenzie(2000)는 연구 단계를 교정 프로그램과 재범 간 상관성 혹은 관련성만 살펴보는 연구(1~2단계), 교정 프로그램을 실행한 집단과 그렇지 않은 집단으로 나누어 비교는 하되 엄격한 통제를 하지 않은 연구(3~4단계), 랜덤통제실험으로 실험집단과 비교집단을 엄격하게 나누어 효과를 분석한 연구(5단계)로 구분하고, 최소 3단계 이상 연구를 검토하여 ‘무엇이 작동하는지’ 혹은 ‘작동하지 않는지’를 따져봐야 한다고 주장하였다. 그는 교정정책의 상당수가 효과에 대한 과학적 증거나 논거도 없이 결정된다는 점을 지적하면서 증거기반 교정정책은 선택이 아닌 필수라는 점을 강조하였다(Mackenzie, 2000: 469).

Welsh와 Farrington(2001) 역시 범죄예방 분야에서 증거기반 관점의 필요성을 주장하였다. 이들은 캠벨 재단이 시작한 범죄예방의 효과성 평가 방식을 따라서 실험 혹은 준실험(experiments and quasi-experiments) 설계를 통한 연구가 필요하다고 보았다. 이러한 실천을 위해서는 정치적 고려와 판단이 필요하다는 사실도 함께 제시하였다.

이처럼 증거기반 형사정책은 정책결정자나 실무자의 경험과 견해에 근거하여 시행되었던 정책이나 프로그램을 과학적이고 합리적 과정을 통해 생산한 경험적 증거에 기반하여 종합적으로 검토하고 평가하지는 학문적 대응이자 노력이다(Cullen and Jonson, 2016). 사실 형사정책이 범죄통계, 경찰통계, 설문조사, 심층면접 등 다양한 경험적 증거를 통해 시행되어야 한다는 주장은 오랫동안 학술공동체에 내재한 일종의 ‘교의’였다. 그런데도 증거기반 형사정책이 많은 관심을 받게 된 이유 중 하나는 증거기반 형사정책이 학계와 실무의 관계를 재정립할 수 있는 계기로 인식되었기 때문이다. 정책실무자는 자신이 담당하는 업무의 개선보다 업무의 안정에 더 관심을 가지는 반면, 연구자는 실무자가 쉽게 이해하고 적용할 수 있는 형태로 연구성과를 제시하는데 인색한 편이다(박철현, 2014). 양측이 각자 원하는 것을 추구하는 동안 연구성과가 형사정책으로 반영되거나 구현되기 어려운 상황에서 증거기반 형사정책은 학문과 정책을 매개하기 위한 ‘화해’의 슬로건으로 매우 적합한 것이었다.

다른 한편으로 증거기반 형사정책은 범죄예방 조치와 관련한 새로운 정책이나 법률안을 입안하고 결정할 때 기본권침해를 수반할 수밖에 없는 국가형벌권 투입으로 얻을 수 있는 효과가 무엇인지 즉, ‘사전 입법 영향 평가’와 ‘사후 입법 성과 평가’를 통해 그 실효성과 파급 영향에 대한 근거를 얻을 수 있다는 점에서 매력적이다. 입법자에게 과학

적 근거는 정책 결정의 정당성과 책임성을 확보하고, 비효율성을 개선하여 정책의 합리성을 높이며, 정책 결정 과정에서 발생하는 이념적·정치적 갈등을 줄이는 기능을 담당할 수 있다(하태훈, 2022: 21). 이러한 측면에서 증거기반 형사정책은 경험적 증거를 중시하는 규범적 관점의 학자들에게도 설득력 있는 개념이라고 볼 수 있다.

앞서 살펴본 것처럼 국내에서 형사미성년자 연령 문제는 오랫동안 논란의 중심에 있다. 그러나 이쉽게도 국내에서는 지금까지 살펴본 엄격한 방법론에 기초한 증거기반 연구를 찾아볼 수 없다. 따라서 이 논문에서는 형사미성년자 연령과 관련하여 미국에서 증거기반의 관점으로 수행된 소년범죄자의 책임능력에 관한 연구, 그리고 소년범죄자의 형사처벌에 관한 연구를 상세하게 살펴보고자 한다. 두 사안은 형사미성년자 연령 논란과 관련하여 그 성격에 다소 차이가 있다. 소년범죄자 책임능력 문제가 형사미성년자 연령을 인정하는 ‘전제(premise)’에 관한 검토라고 한다면, 소년범죄자 형사처벌 강화는 형사미성년자 연령하향 같은 제도 도입이 가져올 예상되는 ‘결과(outcome)’에 관한 검토이다. 그러나 두 사안과 관련한 연구 사례들은 엄격한 방법론을 적용하여 최선의 결과를 도출한다는 증거기반 원칙과 관점을 따르려고 했다는 점에서는 공통적이다. 따라서 두 분야에 대한 체계적 검토를 통해 그 결과와 증거를 살펴보고 합의를 도출해 보고자 한다.<sup>4)</sup>

4) 이 논문에서 다루고 있는 미국의 “책임능력관련 연구”와 “형사이송관련 연구”가 단순히 병렬적 관계가 아니라 서열적 관계일 수 있다는 논평자의 지적이 있었다. 그에 따르면, “형사이송관련 연구”가 “책임능력관련 연구”보다 주가 되어야 한다는 것이다. 뇌과학 연구는 형사미성년자 제도의 필요성에 대한 좋은 증거가 될 수는 있지만 논문이 제기하고 있는 형사미성년자 연령 기준에 대한 증거는 아닐 수 있고, 1살 차이 역시 뇌과학적으로 의미가 크다고 보기 어려울뿐더러 1살 더 낮추자는 주장이 소년을 성인과 똑같이 처우하자는 것은 아니라는 주장이다. 이러한 문제제기가 타당한 측면은 있으나 국내에서 근거로 제시되고 있는 연구 중 하나가 뇌과학 분야이기 때문에 이 부분도 비중 있게 다루지 않을 수 없었다. 다만, 논평자의 의견처럼 저자들 역시 자료를 수집하고 논문을 작성하면서 뇌과학 분야의 연구성과보다 형사이송관련 연구가 국내 형사미성년자 연령 문제에 조금 더 가까운 논거가 될 수 있다고 여겼으나 두 분야의 연구성과를 서로 다른 비중으로 다루지는 않았다.

## IV. 소년사법정책에 관한 증거기반 연구: 미국의 사례를 중심으로

### 1. 책임능력(doli capax) 관련 연구

범죄학에서 연령과 관련한 가장 대표적인 논의는 ‘연령-범죄 곡선(age-crime curve)’이다(Gottfredson and Hirschi, 1990). 연령-범죄 곡선이 형사미성년 논의에서 시사하는 바는 청소년기와 성인기는 분명한 차이가 있다는 것이다. Wilson과 Herrnstein(1985)은 반사회적 행동과 관련된 인구사회적 변수와 상황적 변수를 통제한 후에도 연령이 여전히 예측변수로서 영향을 미친다는 것을 발견했다. Shavit와 Rattner(1988)은 이런 결과에 대해 연령과 반사회적 행동 사이에는 이미 알려진 사회학적 변수로는 설명되지 않는 변량이 있다고 보았다. 호르몬이나 뇌성숙 등 생물학적 요인에 관한 연구가 축적되면 연령과 반사회적 행동 간 설명되지 않는 나머지 변량도 설명할 수 있다는 주장이다(Walsh and Beaver, 2009).

이후 연령에 따른 변화에 관한 심리학적 연구들이 먼저 발표되었다. 일례로 Cauffman과 Steinberg(2000)는 의사결정능력이 연령에 따라 실제로 다른지 분석하였다. 이들은 청소년 1,015명을 대상으로 설문조사를 수행하여 8학년, 10학년, 12학년, 청년, 성인으로 구분해 집단별로 의사결정능력이 다른지를 비교하였다. ‘성숙한 판단능력(maturity of judgment)’을 분석하기 위해 이들은 책임능력(responsibility), 인지능력(perspective), 조절능력(temperance) 등 세 가지 심리사회적 요인을 측정하였다. Cauffman과 Steinberg(2000: 756)는 16세~19세까지 심리사회적 성숙이 가장 두드러지며, 19세에 도달하면 성숙한 판단능력이 더 이상 증가하지 않고, 21세를 넘어서면 이러한 능력은 점차 안정되는 경향을 발견하였다. 심리사회적 요인을 각각 분석한 결과와 합쳐서 분석한 결과 모두 소년과 성인 간 성숙한 판단능력에 차이가 있는 것으로 나타났고, 나이가 가장 어린 8학년일수록 성숙한 판단능력이 가장 저조하게 나타났다. 이들은 비록 소년의 인지능력이 성인과 유사한 수준을 보일 수는 있으나, 다른 능력에서는 소년이 성인에 비해 낮은 수준을 보이므로 법적으로나 정책적으로 소년을 성인과 동등하게 고려해서는 안 된다고 주장하였다.

소년범이 법정에 설 수 있는 역량(competence to stand trial)을 살펴본 Cooper(1997)

의 연구, Baerger와 동료들(2003) 연구도 주목할 만하다. Cooper는 법정에서 소년범의 이해능력 혹은 법정 역량이 연령에 따라 다를 것으로 보았다. 이를 위해 처음 재판을 경험한 11세~16세 소년범 112명을 대상으로 역량검사를 실시하였다. 검사결과, 사전조사에서 2명을 제외하고 모든 소년범이 기준 이하의 점수를 받았다. 이는 법정에 처음 선 소년범의 역량이 전반적으로 충분하지 않음을 시사한다. 특히, 13세 이하의 소년범과 14세 이상의 소년범의 격차가 현저하다는 결과가 나타나 연령에 따라 소년범의 법정 역량에 차이가 있음을 보여주었다. Berger와 동료들도 유사한 주제로 비교연구를 수행하였다. 이들은 법정 역량이 충분하지 않다고 판단되는 132명의 소년범과 법정 역량이 충분하다고 판단되는 473명의 소년범을 통계적으로 비교하였다. 그 결과, 12세 이하의 소년범과 15세 이상의 소년범 간의 차이가 두드러지게 나타났다. 즉, 12세 이하일수록 법정 역량이 불충분한 집단으로, 15세 이상일수록 법정 역량이 충분한 집단으로 분류될 수 있다는 결과를 제시함으로써 연령에 따른 법정 역량의 차이가 존재함을 보여주었다. 이러한 연구는 청소년이 성인과 달리 법적 책임을 질 수 있는지에 관한 심리학적 증거로 제시되었다. 그러나 보다 확실한 증거를 추구하는 강성과학(hard science)의 연구자들은 불완전한 자기보고식 응답에 기반한 연구보다는 소년의 책임능력을 입증할 만한 신경생물학적 증거를 찾기 시작했다.

실제로 2005년 미국 대법원은 신경생물학적 증거를 채택하여 소년의 ‘제한된 책임능력’을 인정하였다.<sup>5)</sup> 미국 대법원은 소년과 성인의 일반적 차이를 규명하는 세 가지 핵심 이유로 소년은 성인에 비해 성숙한 판단을 내리기 어렵다는 점, 소년은 성인보다 부정적 영향 및 다른 외부 영향에 취약하거나 민감하다는 점, 소년이 성인보다 변화(개선) 가능성이 크다는 점을 제시하였다(Venturelli, 2021: 1161). 이 판결은 소년의 형사책임이 과학적 근거를 통해 감경될 수 있음을 인정하면서도 완전한 책임능력 부재를 인정한 것은 아니었다. 2010년 *Graham vs. Florida* 판례에서도 소년과 성인 사이에 명백한 차이가 있음을 심리학 및 뇌과학적 증거를 통해 인정하였다(이인영, 2015: 444-445). 미국 대법원이 과학적 증거를 들어 소년의 책임능력이 성인에 미치지 못한다고 인정한 것은 뇌과학 연구가 신뢰할 만하다는 사실을 반영한다. 그렇다면, 어떠한 증거가 청소년과 성인의 책임능력 차이를 규명한 것인지 증거기반의 관점에서 관련 연구를 살펴보자.

5) *Roper v. Simmons*, 543 U.S. 551, 569-74 (2005)

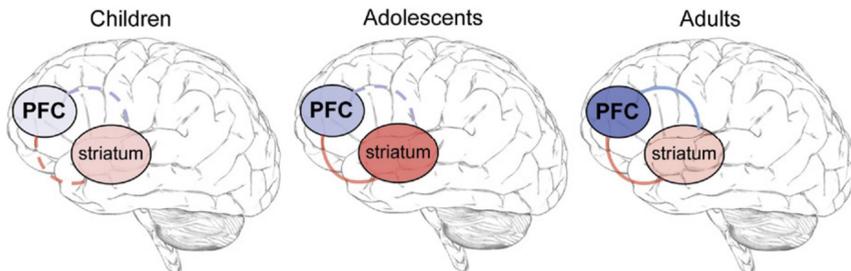
먼저, Casey와 동료들(2008)은 인간과 동물의 뇌를 연구한 결과를 종합적으로 검토하여 청소년 비행이나 위험 행동의 생물학적 설명을 제시하였다. 이들은 기존의 연구를 바탕으로 청소년이 위험한 행동에 더 많이 가담하는 원인은 뇌 부위의 성숙(maturation)이 균등하지 않기 때문이라는 가설을 제시했다. 이들은 충동성과 위험 추구를 담당하는 뇌 부위에서 발생하는 신경생물학적 발달궤적의 차이가 청소년기 비행성을 유발한다는 ‘성숙 불균형 가설(maturation imbalance model)’을 제시하면서, MRI, DTI, fMRI를 활용한 인간과 동물의 뇌에 대한 신경생물학적 근거로 자신들의 가설을 뒷받침하였다 (Casey et al., 2008: 2).

뇌과학에서는 전전두피질(prefrontal cortex, PFC)이 성숙의 최종단계에서 완성된다고 본다. 초기 뇌과학자들은 이 부위(PFC)의 미성숙함 때문에 청소년이 위험한 행동을 선택한다고 생각하였다. 그러나 PFC의 미성숙 때문에 감정이입이 예민해지고 비행에 가담하는 것이라면, 그보다 미성숙한 아동에게서 더 선명한 패턴이 관찰되어야 한다. Casey와 동료들은 청소년이 성인뿐 아니라 아동과도 다른 양상을 보이기 때문에 PFC 미성숙만으로 청소년의 비행을 설명할 수는 없다고 보았다. 이들은 청소년의 충동성과 위험 추구 성향이 뇌의 서로 다른 인지 및 신경 발달과정에 의한 것이라고 보았다. 충동 조절 능력 혹은 만족지연 능력은 인지발달을 통해 형성되는데, 인지발달은 지능의 차이가 아니라 정보처리의 속도와 효율에 기초한 것이다. 즉, 목표한 것을 성취하기 위해 다른 부적절한 생각이나 행동을 억제하는 능력은 인지발달을 통해 이루어진다. 이들은 청소년 충동성이 PFC 발달과 관련이 있으며, 위험행동 추구는 변연계의 발달과 관련이 있다는 사실을 토대로 두 부위 사이의 성숙 불균형에 주목하였다. Casey와 동료들이 그들의 주장을 피력하면서 증거기반으로 제시한 근거가 Galvan과 동료들(2006; 2007), Sowell과 동료들(2003), Liston과 동료들(2006), 그리고 Ernst와 동료들(2005)의 연구이다.

Galvan과 동료들은 청소년기 PFC의 불충분한 발달로 제어 능력이 완전히 기능하지 못할 수 있다고 주장했다(Galvan et al., 2007: 12). 이들은 청소년기 위험한 선택을 하거나 정서적 정보를 처리할 때 허부피질에서 활동성이 높게 관측된다는 결과도 제시하였다(Galvan et al., 2006: 6885). Sowell 등(2003)은 MRI를 통해 전전두피질(PFC)와 기저핵(basal ganglia)의 부피를 측정하여 인지능력 발달은 PFC 같은 피질뿐 아니라 기

저해와 같은 하부피질 발달도 영향을 미친다고 주장하였다. 뇌의 각 부위가 독립적으로 발달하기보다 나이에 따라 부위 간 연결성(connectivity) 강화가 인지능력 향상과 관련이 있다는 사실도 확인했다(Cascio et al., 2007: 215). 부위 간 연결성 강화는 백질신경로(white matter tract) 발달을 통해 일어나는데, 이 경로를 통해 정보가 다른 부위로 이동하기 때문에 뇌 부위 간 연결성 강화는 인지능력의 향상과 관련이 있다(Casey et al., 2008: 4). Liston과 동료들(2006)도 전전두피질과 기저핵 간의 연결성이 충동조절(impulse control)과 관계가 있다는 사실을 밝혔다. fMRI를 활용한 Ernst와 동료들(2005) 연구는 청소년이 아동이나 성인에 비해 보상에 대한 과도한 즉좌핵 반응이 관찰된다고 보고하였다. 아동의 경우에는 전전두엽 통제를 담당하는 부위의 반응이 성인에 비해 훨씬 떨어지는 것으로 나타났다. 전전두엽 부위 및 하부피질의 발달은 연령에 따라 서로 다른 발달궤적을 보이며, 청소년기 과민한 즉좌핵 반응이 충동적이고 위험한 행동에 기담하는 것과 밀접한 상관성이 있다. 결론적으로 청소년기는 전전두피질의 미성숙 때문에 보상에 대한 충분한 통제력을 행사하지 못하는 동시에 변연계의 영향력이 더 강성하여 정서 정보에 대한 합리적 반응은 제한되고, 보상 중시적 결정은 취약해진다(Casey et al., 2008: 7). Casey와 동료들은 이 모든 연구를 ‘성숙 불균형 가설’을 지지하는 증거로 사용하였다. 청소년이 감정적 상황에 처할 경우 변연계 작용이 전전두엽 통제기능을 앞서는데, 이런 상황에서는 자신의 행동이 위험하다는 것을 알고 있더라도 위험한 행동을 선택하게 된다는 것이다(Casey et al., 2008: 7).

[그림 1] 연령별 피질 부위와 하부피질 간 연결성 시각화



출처: Casey & Jones, 2010. 2면.

Dreyfuss와 동료들(2014)도 fMRI 연구를 통해 청소년이 아동이나 성인에 비해 사회적·정서적 신호에 더 민감하게 반응하며, 자기조절능력은 제한될 것이라는 가설을 제시하였다. 6세부터 27세까지 57명(남성 30명, 여성 27명) 대상으로 fMRI를 사용하여 go/no-go 테스트를 진행하였다. 참여자들은 공포, 행복, 평온 세 가지로 구성된 화면에 나타난 얼굴표정을 보고 ‘버튼을 누르거나(go)’ ‘누르지 않도록(no-go)’ 안내를 받았다. 이 테스트는 참여자가 공포에 질린 얼굴표정과 평온한 얼굴표정에서 ‘no-go’를 얼마나 잘못 누르는지 파악하려는 실험이다. 참여자에게 얼굴표정을 무작위 방식으로 제시하여 회당 48번(36번의 go, 12번의 no-go)씩 총 2회를 누르도록 실험한 결과 연령에 따른 오검출률<sup>6)</sup>이 유의미하였는데( $F=8.58, p<0.001$ ), 청소년이 아동과 성인보다 공포에 대한 오검출율이 높았다. 오검출율과 좌측 안와전두피질의 신호 반응에서는 남자 청소년의 활동성이 두드러진 반면, 여자 청소년은 회피관련 행동 조절을 담당하는 내측 전전두피질 반응에서 활동성이 두드러졌다. 이들은 ROI 분석을 통해 좌측 안와전두피질과 내측 전전두피질에서 나타난 패턴이 청소년기에만 특별히 나타나는 효과(adolescent-specific effects)라는 것을 발견하였다. Dreyfuss와 동료들은 보상신호가 위협감수나 범죄행동 등 청소년의 충동적인 의사결정을 야기한다는 기존 연구(Somerville et al., 2011)를 바탕으로 위협 신호(threat cue)가 충동조절에 미치는 영향도 살펴보았다. 이들은 청소년기 비합리적 행동 특성이 청소년기 동안 내측 전전두피질 부위의 발달과 하부피질 간의 연결성이 충분히 발달하지 않는 데서 비롯된다고 보았다(Liston et al., 2006; Somerville et al., 2011). 내측 전전두피질이 행동의 조절을 담당하는 만큼 전전두피질과 변연계 연결은 인지조절 능력에서 중요하다. 그러므로 양자 간 약한 연결성은 인지조절 능력을 감소시키며, 특히 감정적으로 고조된 상황에서 더 약해진다(Cohen and Thomas, 2013). 이들은 아동이나 성인과 달리 청소년이 위협 신호에 대한 허위경고(false alarm)에 두드러지게 반응하는 이유 역시 안와전두피질과 내측 전전두피질 부위 활성화 증가와 관련이 있다고 주장하였다. 이 부위는 정서와 행동반응을 조절하며 특히 위협과 관련한 자극을 다룬다. 관련 없는 정보를 무시하는 능력이 청소년기 가장 활성화된다는 사실도 선행연구(Casey et al., 2000)와 일치하였다(Dreyfuss et al., 2014: 225). 청소년기는 정서 관련 정보가 긍정적일 때보다 부정적일 때 더 쉽게 영향을 받는다는 연구(Cohen and

6) no-go 상황이면 버튼을 누르지 않아야 하지만, 실제 테스트에서 버튼을 누른 정도를 나타낸다.

Thomas, 2013)에 근거하여, 연구자들은 청소년기 변연계 회로 변화가 정서적 신호에 대한 높은 민감성을 부여하여 잠재적 위협 신호로부터 후퇴하기보다 충동적으로 반응하는 결과를 야기한다고 해석하였다.

지금까지 책임능력에 대한 뇌과학 연구와 그 증거를 살펴보았다(표 1. 참조). 결론적으로, 청소년기의 충동성과 위험추구 성향은 연령에 따른 미성숙과 관련이 있다. Casey와 동료들이 주장한 ‘성숙 불균형 가설’처럼 변연계의 이른 성숙과 전전두엽 부위의 늦은 성숙이 상호작용하여 청소년기의 비행 성향에 기여한다. 그러나 Epstein(2007)은 MRI와 같은 뇌의 촬영기술을 통해 제시한 증거는 ‘한순간만 포착한 것(snapshot)’에 다름없고, 그러한 증거는 상관관계를 나타내는 것이지 인과관계를 의미하는 것은 아니므로 청소년의 인지능력이 제한되어 있다고 볼 수만은 없다고 반박하였다. 그는 청소년기 혼란(turmoil)은 오히려 문화적 요인에서 기인한다고 주장하였다. 뇌과학 증거들이 청소년이 성인과 다르다는 특성에 관한 상관성 증거로 유효할 수 있으나, 인과적 추론에는 여전히 무리가 있다는 비판이다.

〈표 1〉 책임능력 증거를 제시한 연구결과 요약

저자 (년도)	표본 (나이)	측정방식 (시험)	뇌 부위	결과 요약
Sowell과 동료들 (2003)	일반인 176명 (7~87세)	MRI와 피질매칭 알고리즘 (없음)	측면반구와 뇌 반구 간 영역	<ul style="list-style-type: none"> <li>회백질 밀도에서 유의미한 비선형 감소가 배측 전두엽과 두정엽의 연합피질에서 관측됨</li> <li>회백질 밀도 내의 연령과 관련된 패턴 변화가 좌반구 후외측두피질에서 역전됨(이 부위에서 회백질 밀도는 약 30살까지도 증가함)</li> </ul>
Liston과 동료들 (2006)	대상자 21명 (7~31세)	DTI (go/no-go 시험)	전두선조체 피질척수로	<ul style="list-style-type: none"> <li>전두선조체의 급격한 확산성은 연령이나 정확도와는 독립적으로 빠른 반응시간을 예측하며, 이러한 상관관계는 더 높은 통제력이 요구되는 테스트에서 더 강력해짐</li> <li>속도시험의 경우, 성인들의 정확성이 훨씬 높았는데, 정확성은 전두선조체의 확산성과 밀접한 상관관계가 있고, 피질척수로와는 연관이 없었음</li> </ul>
Galvan과 동료들 (2006)	대상자 37명 (7~29세)	사건 관련 fMRI (두 선택지 지연반응 시험)	전전두엽 측좌핵 안와전두피질	<ul style="list-style-type: none"> <li>아동이나 성인의 전전두엽 활동성에 비해 측좌핵 활동성이 높게 나타났으며, 이는 두 부위가 서로 다른 발달 시기를 겪는다는 것을 뜻함</li> <li>청소년의 측좌핵 활동과 보상가치에 대한 활동성과 민감성 수준이 성인의 것과 거의 유사한 반면, 안와전두피질의 활동성은 아동에 더 가까웠고, 더 분산된 활동성을 보였음</li> </ul>

저자 (년도)	표본 (나이)	측정방식 (시험)	뇌 부위	결과 요약
Galvan과 동료들 (2007)	대상자 37명 (7~29세)	fMRI (무기명 자기보고식 위험행동 평가)	측좌핵	<ul style="list-style-type: none"> <li>• 측좌핵 활동성과 위험행동 가담 가능성 간의 정적인 상관관계가 전반에 걸쳐 나타난 반면, 충동성 평가는 측좌핵 활동성과는 관련이 없고 연령과 관련이 있었음</li> <li>• 청소년기 동안, 일부는 단순히 충동성의 변화 때문이라기보다 위험행동에 가담하는 성향의 변동성과 발달 변화가 맞물려 위험행동에 특별히 더 가담하게 됨</li> </ul>
Cohen과 Thomas (2013)	대상자 100명 (11~25세)	자기보고식 조사 (go/no-go 시험)	-	<ul style="list-style-type: none"> <li>• 연령차이에 따른 억제시험상의 정확성을 보면 다른 시기와 비교할 때 이른 청소년기에 부정적인 감정동요에 의해 억제반응이 더 쉽게 방해받는 것이 발견됨</li> </ul>
Dreyfuss와 동료들 (2014)	대상자 80명 (6~27세)	fMRI (go/no-go 시험)	전전두엽 피질 변연계 피질 회로	<ul style="list-style-type: none"> <li>• 청소년기는 사회적 그리고 정서적 신호에 대한 민감도가 높아진 시기로, 위협신호 앞에서 행동 조절능력이 제한됨</li> <li>• 청소년기의 충동적인 행동은 위협신호의 존재가 보상신호로 작동하여 반응을 철회하거나 보류하지 않고, 충동적으로 반응하는 경향을 보임. 이런 경향성은 반응하지 않도록 지시한 상황에도 동일했음</li> </ul>

## 2. 형사이송(criminal transfer) 관련 연구

미국은 국친사상(parens patriae)을 바탕으로 소년범죄자에 대해 온정주의를 펼쳐왔으나, 1980년대 들어 여론이 소년범죄의 심각성에 우려를 나타내면서 강경한 처벌로 선회했다(박미랑, 2014). 특히 형사이송제도(juvenile transfer 혹은 waiver) 개정에서 처벌 강화 의지가 두드러지게 표출되었다. 형사법원으로 이송되어 재판을 받거나 형사처벌을 받을 수 있는 기준을 확대하는 등 법률개정을 통해 많은 소년범이 소년법원이 아닌 형사법원에서 무거운 형벌을 선고받고 성인교도소에 수용되었다. 소년범에 대한 형사이송은 억제이론(deterrence theory) 입장에서 소년 역시 성인과 마찬가지로 형벌에 따른 범죄 억제가 발생할 것이라는 논리에서 비롯된 제도이다.

소년범에 대한 엄한 처벌이 소년범죄율 감소에 긍정적인 영향을 미칠 수 있다는 근거로는 Levitt(1998)의 연구가 있다. 그는 형사이송제도와 같이 형사처벌의 범위를 확대할 때 전체 소년범죄 발생률이 감소하는지 실증적으로 분석하였다. 1978년부터 1993년까지

지 주 단위 패널 데이터를 분석한 결과, 소년범죄에 대한 강력한 처벌이 소년범죄율의 감소와 상당한 관련이 있는 것으로 나타났다. 그에 따르면, 1978년부터 1993년 기간 동안 소년에 대한 처벌율은 20% 감소한 반면, 성인에 대한 처벌율은 60% 증가했다. 같은 기간 소년범죄는 폭력범죄 107%, 재산범죄 7%가 증가한 반면, 성인범죄는 폭력범죄 52%, 재산범죄 19%가 증가했다. 만일 소년에 대한 처벌율이 성인만큼 높았다면, 같은 기간 폭력범죄는 74%, 재산범죄는 2% 수준에서만 증가했을 것이라는 게 그의 주장이다(Levitt, 1998: 1181). 다만, 소년법에 대한 엄격한 처벌이 성인이 된 이후 범죄까지 미칠 영향에 대해서는 불확실하다고 보았다. 그는 비용-편익(cost-benefit) 측면에서 볼 때, 소년을 엄격하게 처벌하는 것이 반드시 효율적인가에 대해서도 의문을 제기했다. Levitt의 연구는 형사이송과 같은 강력한 처벌 정책이 소년범죄를 억제하는데 효과가 있을 수 있다는 결과를 제시하고 있으나, 형사이송제도 자체를 평가하기보다는 엄벌주의 정책 전반에 대한 효과를 살펴본 것이고 실험설계에 기반하지 않은 결과라는 점에서 비판받을 수 있다.

이후 강경한 소년사법제도의 효과를 평가하는 연구가 하나둘씩 등장하면서 형사이송의 확대가 재범률 감소에 긍정적인 영향을 미치지 못하다는 비판이 제기되었다(Fagan et al., 2004; Winner et al., 1997; Fowler and Kurlychek, 2018; Zane et al., 2016). 특정 범죄유형이나 범죄의 심각성에 따라 일부 효과가 있다고 주장되기도 했으나, 대체로 엄한 처벌이 부과되었을 때 소년의 재범률이 높거나 재범까지 걸린 기간이 짧아지는 역효과가 발견되었다. 미국의 형사이송제도에 관한 증거기반 연구는 국내 형사미성년자 연령 논의와 관련하여 시사점을 제공한다(박미량, 2014; 박선영, 2019). 여기서는 증거기반에 기초한 4개의 관련 연구를 살펴보고자 한다.

첫째, Winner와 동료들(1997) 연구는 플로리다 보건부에서 수집한 Client Information System(CIS) 데이터를 활용하였다. 1987년 1년간 형사이송된 사건 3,142개를 모집단으로 하여 형사미이송집단과 매칭(matching)을 통해 각각 2,700명씩 표본을 선정하였다. 종속변수는 재체포율과 체포까지 걸린 기간이며, 재체포 자료는 플로리다 법무부에서 1994년 11월까지 수집한 것을 사용하였다. 분석 결과, 형사미이송집단은 1,128명(42%), 형사미이송집단은 1,163(43%)명이 재체포되었다. 재체포율 차이는 통계적으로 유의미하지 않았으나, 경범죄는 형사미이송집단이 형사미이송집단에 비해 재체

포율이 높았고( $z=-2.79$ ,  $p=0.005$ ), 중범죄는 형사미이송집단이 형사이송집단에 비해 재체포율이 더 높았다( $z=2.53$ ,  $p=0.011$ ). 범죄유형별로는 재산중범죄에서 형사미이송집단이 재체포율이 더 높았고( $z=5.47$ ,  $p<0.001$ ), 다른 범죄유형은 형사이송집단이 재체포율이 높았다(Winner et al., 1997: 552).

로짓분석에서는 형사이송이 재체포에 미치는 영향력이 유의하였다(Wald=4.378,  $p=0.036$ ). 재산 중범죄는 재체포 가능성을 높이는 것으로 나타났지만(Wald=16.533,  $p<0.001$ ), 형사이송과 재산 중범죄의 상호작용은 재체포 가능성을 낮추는 것으로 나타났다(Wald=17.658,  $p<0.001$ ). 생존분석에서는 형사이송이 재체포 가능성을 높이는 것으로 나타났다. 전체범죄를 기준으로 형사이송집단의 재체포 기간이 형사미이송집단보다 짧게 나타난 것이다((O-E)T=45.26,  $p=0.058$ ). 다만, 형사이송집단 생존율은 빠르게 떨어지다가 1,500일 시점에서 형사미이송집단 생존율과 교차하며 이후 약하게 역전되는 양상을 보였다. 중범죄를 기준으로 두 집단 간 차이는 유의하지 않았으나, 경범죄 기준으로는 형사이송집단이((O-E)T=32.48,  $p=0.002$ ), 재산 중범죄 기준으로는 형사미이송집단이((O-E)T=-47.46,  $p=0.001$ ) 재체포까지 걸린 기간이 더 짧았다. 결론적으로 Winner와 동료들(1997: 559)은 재산 중범죄를 제외하고 형사이송제도가 재범을 억제하기보다 오히려 재범을 증가시키는 부적 효과가 있다고 주장하였다. 따라서 형사이송을 고려할 때 일률적 적용보다는 효과가 있는 범죄유형에 차별적으로 적용할 것을 제안하였다.

둘째, Fagan과 동료들(2004)은 소년범을 소년법원이 아닌 형사법원에서 처벌하는 것이 공공의 안전을 확보하는데 긍정적 효과가 있는지 연구하였다. 형사이송제도의 효과를 증거기반 연구로 수행하는 것이 현실적으로 상당한 한계가 있으나 Fagan 등은 서로 다른 사법체계에서 유사한 죄질의 소년범을 다룬 결과를 비교한 준실험설계 방식으로 접근하였다. 표본편향의 문제를 해소하기 위해 ‘대도시 통계지구(Metropolitan Statistical Area)’에서 유사한 배경을 가진 뉴욕주와 뉴저지주 카운티에서 각각 표본을 추출하여 매칭을 시도하였다(Fagan et al., 2004: 18). 모집단은 1992~1993년 동안 강도, 가중폭행, 절도를 저지르고 소년법정 혹은 형사법정으로 기소된 청소년(15세~16세)으로 2,382명을 표본으로 선정했다. 뉴욕주는 강도, 가중폭행, 절도죄로 기소된 소년범을 예외적으로 형사법정에서 처리하며, 뉴저지주는 동일한 범죄를 저지른 소년범을 소년법정에서 처

리한다. 소년법정 사건은 뉴저지 카운티에서, 형사법정 사건은 뉴욕시 카운티에서 추출하였다.

Fagan과 동료들은 법정(소년법원 vs. 형사법원)에 따른 처벌 수위의 차이를 확인하기 위해 유죄판결 비율과 유죄판결에 따른 제재 유형을 비교하였다. 분석 결과, 뉴저지주(소년법원)에 비해 뉴욕주(형사법원)가 구금형 비율이 높았다(10.7% vs. 32.6%). 비록 뉴저지주가 뉴욕주에 비해 유죄판결 비율이 높았지만(61.3% vs. 52.0%), 전반적인 처벌에 있어 뉴욕주보다 관대한 양상을 보였다. 재체포율을 살펴보면, 전체적으로 뉴저지주가 뉴욕주보다 높았지만(72.4% vs. 69%), 마약을 제외하고 폭력, 재산, 무기소지 등 다른 범죄유형에서 뉴욕주가 더 높았다.

생존분석에서 법정유형의 효과가 모든 유형의 재체포에서 유의미하지 않았으나 폭행 및 재산범죄는 형사법정(뉴욕주)에서, 마약범죄는 소년법정(뉴저지주)에서 재체포 가능성이 높게 나타났다. 형사법정을 거친 소년은 재수감률 또한 소년법정보다 25.5% 높았다. Fagan과 동료들은 정책입안자 의도와 다르게 형사이송제도의 재범억제에 대한 증거는 발견하지 못했다고 주장하였다. 오히려 폭행 및 재산범죄는 형사법정에서 처리할 경우 더 많이, 더 빨리 재범을 저지르는 것으로 나타났다. 다만, 마약범죄는 소년법정에서 처리할 경우 오히려 재범가능성이 높게 나타났다. 이러한 결과는 연령처럼 특정 기준에 따라 일괄 적용하는 제도는 특별억제에 효과적이지 못하다는 것을 보여주는 결과로 해석하였다(Fagan et al., 2004: 66).

셋째, Fowler와 Kurlychek(2018) 연구는 자연실험(natural experiment)에 기반하였다. 코네티컷주가 2010년 1월부터 ‘연령상향법(the Law of Raise the Age)’에 따라 형사이송 연령을 16세에서 17세로 상향함에 따라 Fowler는 새로운 법률 적용 전후 자연적 비교를 수행하였다. 2010년 이전에 16세였던 소년범은 형사법정에서 사건이 처리되었고, 2010년 당시 16세였던 소년범은 소년법정에서 사건이 처리되었기 때문에 두 집단을 비교함으로써 형사이송 연령 상향이 재범에 미치는 영향을 확인할 수 있다. 연구자들은 소년법정과 성인법정 중 어떤 관할에서 소년범을 처리할 때 재범이 감소하는지를 검증하였다. 자료는 코네티컷 사법부로부터 2009~2010년 기소된 16세 소년의 체포기록, 범죄경력, 재범자료를 사용하였다. 연도별로 기소된 표본 수는 2009년 1,360명, 2010년 901명이다. 분석은 재범자로 구분된 표본을 모두 포함한 경우(N=7,153), 연도

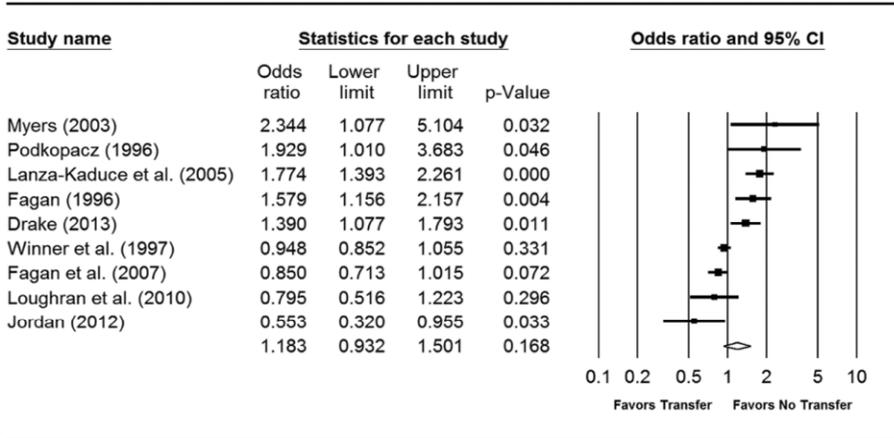
별로 기소된 표본만 대상으로 한 경우(N=2,261)로 나누어 실시하였다(Fowler and Kurlychek, 2018: 269).

분석 결과, 2009년 형사법정에서 처분받은 소년 중 42%가 재체포되었고, 2010년 소년법정에서 처분받은 소년 중에는 26%가 재체포된 것으로 나타났다. 다른 변인을 통제하고 로지스틱 분석을 실시한 결과, 전체 표본과 기소된 표본 모두 소년법정에서 처리될 수록 형사법정에서 처리된 것보다 재체포율이 유의미하게 낮았다. 이들은 결과의 신뢰성을 높이기 위해 민감도 분석을 진행하였는데 결과의 차이는 없었다. Fowler와 Kurlychek(2018)은 16세 소년을 성인처럼 형사법정에서 처리하는 것이 재체포 가능성을 높인다는 발견을 통해 코네티컷주 ‘연령상향법’이 재범억제에 긍정적 영향을 미친다고 보았다.

넷째, Zane과 동료들(2016)은 형사이송제도 연구에 관한 메타분석을 실시하였다. 이들은 600건의 관련 자료를 수집하고, 그중에서 164건이 연구와 관련 있는 것으로 판단하여 1차 추출하였다. 연구결과를 제시하지 않거나 재범을 결과변수로 사용하지 않은 143건은 제외하고, 남은 21개 중에서 비교군이 없는 2건, 실험집단에 소년범죄자가 없는 2건, 재범이 종속변수가 아닌 2건, 표본이 겹치는 6건을 제외한 9건을 최종 분석하였다(Zane et al., 2016: 907).

분석 결과, 5개 연구는 형사이송제도가 소년법정에서 사건을 처리했을 때보다 재범이 증가했고, 3개 연구는 통계적으로 유의미하지 않은 것으로 나타났다. 1개 연구만 유일하게 형사이송이 재범을 감소시키는 것으로 나타났다. 모든 연구를 종합한 평균 효과크기는 통계적으로 유의미하지 않았다(OR=1.183, p=0.168). 달리 말해, 형사이송이 소년범의 재범에 미치는 영향은 긍정적이지도 부정적이지도 않아서 범죄통제나 재범감소와 같은 정책결과에 아무런 영향을 미치지 못한다는 것이다. 이들은 추가분석을 통해 죄질이 심각한 중범죄 유형, 그리고 중범죄 중에서도 폭력범죄만 따로 분석을 하였다. 그 결과, 형사이송이 중범죄의 재범에 미치는 영향은 3개 연구에서 재범이 증가하였고 1개는 재범이 감소했으나 중범죄 전체에서는 효과가 유의미하지 않았다. 중범죄 중 폭력범죄만 따로 분석한 결과에서는 2개 연구에서 재범이 증가하는 것으로 나타났고 폭력범죄 전체에서도 효과가 유의미하여 형사이송이 폭력범죄의 재범은 억제하지 못한다는 사실을 보여주었다.

[그림 2] Forest Plot 분석 결과표



출처: Zane, Welsh, & Mears, 2016. 910면.

Zane과 동료들은 동질성 검증을 통해 연구 간 차이가 있다는 점을 발견하고 연구설계에 따른 메타분석을 추가로 하였다. 분석 결과, 자연실험 연구(OR=1.124, n.s.)와 단순 매칭 연구(OR=1.284, n.s.)는 통계적으로 유의미하지 않고 이질성도 높게 나타났다. 통계적인 통제를 거쳐 비교한 연구(OR=2.089,  $p < 0.004$ )는 형사이송을 한 경우가 그렇지 않은 경우보다 재범 가능성이 더 높았다. 하지만, 성향점수매칭(OR=0.690,  $p < 0.036$  연구)는 형사이송이 그렇지 않은 경우보다 재범 가능성이 더 낮았다. 결론적으로 이들은 메타분석 결과가 형사이송의 특별억제 효과가 없다는 선행연구 결과와 일치한다고 주장하였다. 이들은 형사이송 연구가 이질적이라는 사실도 발견하였다. 다만, 연구 간 이질성은 표본오차 영향은 아닌 것으로 보이며 형사이송뿐 아니라 다른 처벌도 차별적 효과가 있음을 지적하였다(Zane et al., 2016: 916). 민감도 분석에서 성향점수매칭 방식을 제

7) 성향점수매칭(propensity score matching)은 과학적 연구방법의 기초가 되는 랜덤통제실험(RCT) 적용이 어렵거나 애초에 불가능한 경우 수집된 관측자료를 통계적 모델링을 통해 RCT와 비슷하게 혹은 이에 준하여 추정할 수 있다는 가정을 적용하여 독립변수가 종속변수에 미치는 영향, 즉 처치효과(treatment effect)를 추정하는 기법이다. RCT 연구는 무작위배치(random assignment)를 통해 자기선택 편향을 통제하지만, 서베이와 같은 관측연구에서는 ‘교란변수를 활용해 개체가 실험집단에 배치될 확률의 의미하는 성향점수(propensity score)를 추정한 후 이를 모형화하여 무작위배치와 동등하거나 유사한 상황을 창출하는 통계적 방법이다(백영민·박인서, 2020: 14). 성향점수매칭은 최근 실험설계가 어려운 사회과학에서 정책의 효과성 평가방법으로 많이 활용되고 있다.

거했을 때 전체 효과크기가 증가하는 것으로 나타났는데, 이러한 결과를 토대로 해석하면 형사이송에서 재범이 증가한다는 결론이 도출된다. 그러나 Zane 등은 연구설계의 중요성을 강조하면서 성향점수매칭이 다른 준실험설계와 비교하여 랜덤통제실험에 가장 가깝기 때문에 내적 타당도가 높다는 점을 간과할 수 없다고 보았다(Zane et al., 2016: 917).

〈표 2〉 형사이송 효과를 제시한 연구결과 요약

저자 (년도)	모집단 (연구기간)	처치집단	통제집단	형사이송 방식	연구설계 (추적기간)	결과변수
Winner와 동료들 (1997)	87년에 CIS 데이터에 들어온 모든 비행소년 (1987-1994)	형사이송 소년 2,700명	형사미이송 소년 2,700명	혼재팀 (대부분 소년을 바로 형사법정으로 소 제기)	준실험설계 표본매칭 (7년)	재체포와 재체포까지 걸린 기간
[연구결과 요약]	<ul style="list-style-type: none"> <li>• 재산 중범죄를 제외하고, 전제적으로 형사이송은 재범을 오히려 증가시키는 효과가 나타남</li> <li>• 형사이송은 일률적 적용보다 효과가 있는 범죄유형에 차별적으로 적용할 필요가 있음</li> </ul>					
Fagan과 동료들 (2004)	중강도, 중폭행, 중절도로 체포된 15-16세 소년 (1992-1993)	뉴욕 형사법정에서 다룬 소년 1,321명	뉴저지 형사법정에서 다룬 소년 1,061명	법적 배제	준실험설계 자연실험 (2년)	모든 재체포
[연구결과 요약]	<ul style="list-style-type: none"> <li>• 미약범죄를 제외하고, 형사이송이 재범을 억제한다는 증거를 발견하지 못했으며, 오히려 폭행이나 재산범죄의 경우 재범을 빠르게 그리고 많이 하는 것으로 나타남</li> </ul>					
Folwer & Kurlycheck (2018)	2009년과 2010년에 체포된 모든 16세 소년 (2009-2011)	형사법정에서 다룬 소년 4,050명	소년법정에서 다룬 소년 3,103명	법적 배제	준실험설계 자연실험 (2년)	모든 재체포
[연구결과 요약]	<ul style="list-style-type: none"> <li>• 형사법정에서 사건이 처리된 경우 소년법정인 경우보다 재범가능성이 높다는 결과가 나타남</li> </ul>					
Zane과 동료들 (2016)	모든 형사이송 관련 연구	-	-	-	메타분석	효과성
[연구결과 요약]	<ul style="list-style-type: none"> <li>• 전체적인 결과는 형사이송제도가 재범에 미치는 영향이 통계적으로 유의미하지 않게 나타남</li> <li>• 연구설계에 따른 민감도 분석에서 성향점수매칭 연구에서는 형사이송이 재범가능성을 낮춘다는 결과가 나타남</li> </ul>					

형사이송 관련 연구의 증거는 전체 재범에서는 그 영향력이 유의하지 않거나, 오히려 재범을 증가시키는 것으로 나타났다. 메타연구도 이와 같은 결론을 보여준다(표 2. 참조). 따라서 소년에 대한 엄벌정책으로서 형사이송은 효과적인 정책이라 할 수 없다. 다만, 메타분석에서 성향점수매칭을 활용한 2개의 연구는 형사이송의 재범억제력을 긍정 하는 결과를 제시했다. 엄격한 연구설계의 중요성을 무시할 수 없기에 향후 더 엄밀한 설계를 통해 도출한 결과를 지켜볼 필요가 있다.

## V. 증거기반 연구의 한계와 전망

이 논문은 정부 주도로 진행되고 있는 형사미성년자 연령 하향 문제를 둘러싼 논의에서 특정 사건에 대한 언론보도나 일부 정치인의 주장이 아니라 증거기반의 연구성과가 객관적이고 합리적인 정책 의사결정에 긍정적 영향을 미칠 것이라는 가정 하에 우리나라보다 앞서 엄벌적 소년사법정책을 시행한 미국의 증거기반 연구사례를 분석하였다. 이를 위해 소년범죄 연령 하향을 둘러싼 찬반 양측의 입장을 살펴보고, 정책의 과학화 및 합리화를 추구하는 증거기반의 개념과 필요성에 대해 논의한 후, 증거기반 관점에서 소년의 책임능력 및 형사이송제도의 효과성에 관한 연구를 구체적으로 검토하였다. 주요 내용을 요약하면, 다음과 같다.

첫째, 소년범죄 연령 하향에 대한 논쟁은 두 가지 측면으로 구분할 수 있다. 먼저, 소년범죄 추세와 관련하여 일부에서는 소년범죄가 증가하였고 범행 또한 잔혹해지는 추세이고, 이는 경제성장 및 사회변화에 따른 것이므로 법률도 이러한 변화가 반영되어야 한다고 주장한다. 반대 입장에서는 전반적으로 소년범죄 추세가 감소하고 있으며, 특정 연령에서만 증가하는 경향이 있어 14세 기준의 연령 하향은 설득력이 없다고 주장한다. 다른 논쟁은 미성년자의 책임능력에 대한 것으로, 일부에서는 소년들이 더 빨리 성숙해지고 있어 14세 미만자에게 일률적으로 형사책임을 면제하는 것은 불합리하다고 주장한다. 그들은 소년들이 법에 대한 이해도가 높아졌으며 범죄의 성격과 질에 따라 적절한 처벌이 가능하도록 법을 개정해야 한다고 주장한다. 반대 입장에서는 소년들의 성숙도를 판단하기 어렵다는 점을 지적하며 일정 연령에 도달하지 않은 사람에게 형벌을 부과할 수 없다는 입법의 목적과 논리를 강조한다. 이들은 형법의 책임주의와 소년법의 보호주의가 교차하는 영역에 형사책임능력이 위치하기 때문에 좀 더 타당한 근거를 가지고 세심하게 논의해야 한다고 주장한다. 미국에서 이미 시행된 엄벌주의 소년사법정책의 부정적인 효과를 소개하면서 우리나라도 소년사법이야말로 증거에 기반한 형사정책이 실천되어야 할 분야라는 점을 강조한다.

둘째, 형사미성년 연령하향 문제를 다루는데 있어 증거기반 연구의 필요성을 검토하였다. 보건의학 분야에서 시작된 증거기반은 20세기 후반 사회과학이 시민과 사회에 미치는 영향력의 한계에 대한 타개책으로 관심을 불러일으켰다. 특히, 정책 분야에서 증거

기반은 기존 연구들 가운데 타당한 이론에 근거하여 엄격한 방법론을 적용한 연구를 선별하고 그 결과를 종합하여 실제로 작동하는 제도나 프로그램이 무엇인지 밝힘으로써 합리적 정책 의사결정을 지원하자는 의도에서 비롯되었다. 서구에서 형사정책 분야의 증거기반은 수없이 발표되는 연구들이 현실과 동떨어진 이론적 논의에 그치고 있다는 비판, 그리고 이론이 아닌 직관과 견해에 의존하는 정책이야말로 비효과적이고 비효율적이라는 비판을 극복할 수 있는 대안으로 여겨졌다. 증거기반 형사정책에 대한 기대는 국내에도 전파되어 과학적이고 합리적인 범죄예방정책과 형사입법으로 나아가기 위한 계기로 삼아야 한다는 주장이 힘을 얻으면서 큰 관심을 받았다.

셋째, 형사미성년 연령하향 문제와 관련하여 미국에서 이루어진 증거기반 연구를 청소년 책임능력과 관련한 뇌과학 연구, 그리고 소년범 처벌강화를 위한 형사이송의 효과성 연구로 나누어 그 결과를 검토하였다. 뇌과학 연구결과는 청소년의 미숙한 뇌 발달로 인한 한정된 책임능력을 긍정하는 것으로 나타났다. MRI나 DTI를 활용한 뇌과학적 증거들은 성숙 불균형 가설을 지지하였고, fMRI를 활용한 연구에서도 청소년의 자기조절 능력이 제한된다는 결론이 도출되었다. 전전두엽 부위의 불충분한 발달과 하부피질 부위의 이른 발달로 인한 불균형 상태에서 청소년은 위험추구나 충동적인 행동을 더 많이 한다는 것이다. 다만, 뇌과학 증거가 상관관계 이상의 강한 증거를 제시하지 못한다는 주장(Epstein, 2007)도 간과해서는 안 된다. 형사이송의 효과성을 검증한 연구들은 일반적으로 재범을 억제하는 효과가 통계적으로 유의미하지 않거나 오히려 역효과가 있음을 보여준다. 재산범죄나 마약범죄의 경우 형사이송이 재범억제에 미치는 영향이 일부 나타났다. 이러한 증거가 형사이송을 일반적으로 지지하는 결과로 보기는 어렵다. 메타분석에서도 형사이송의 효과는 전체적으로 볼 때 유의미하지 않았고, 폭력범죄의 경우 형사이송이 재범을 억제하지 못하는 것으로 나타났다. 흥미롭게도 연구설계에 따른 차이 분석에서는 성향점수매칭을 활용한 연구에서 형사이송의 재범억제 효과가 있는 것으로 나타났다. 따라서 더 엄격한 실험설계를 통한 증거를 간과할 수는 없으며, 향후 엄밀하게 설계된 연구를 통해 무엇이 실제인지 밝힐 필요가 있다.

결론적으로 미국의 연구 사례들이 현재 국내에서 논의되고 있는 형사미성년자 연령하향을 당장에 지지한다고 보기에는 무리가 있는 것으로 여겨진다. 그럼에도 뇌과학 분야의 연구결과가 보여주고 있듯이 청소년일수록 성인보다 개선 가능성이 높다는 사실을

고려할 때, 소년보호의 관점에서 처벌보다는 교화가 더 효과적일 것으로 여겨진다. 앞서 살펴본 증거들이 모두 미국의 연구 사례들이기 때문에 한국의 상황에 직접 적용하여 예측하기에는 무리가 있는 것도 사실이다. 그러나 뇌 성숙의 증거는 지역에 따른 영향이 상대적으로 미미한 영역의 증거로 고려할 만한 가치가 충분하다고 여겨진다. 형사이송에 관한 증거들도 미국과 한국의 제도적·법률적 차이를 간과할 수 없기에 국내 형사정책의 직접 증거로 삼기에는 무리가 있을 수 있으나 합리적이고 객관적인 증거기반의 연구를 기획하기 위해서는 참고할 만한 가치가 충분하다고 여겨진다. 지금은 강경한 처벌을 우선으로 선택하기보다 양질의 증거를 축적하고 처벌 이외 대안까지 신중하게 논의할 때라고 말할 수 있다.

다른 한편으로 지식 담론에 간혀 무용론이 제기된 위기의 사회과학과 갈수록 불확실하고 복잡한 현실에서 해결책을 제시해야 하는 정책당국이 대안으로 선택한 증거기반 역시 만병통치약(panacea)이 아니라는 주장에도 귀를 기울일 필요가 있다(한승훈·안혜선, 2021: 303). 증거기반에 관한 다음과 같은 비판은 증거기반 연구의 수행 시 보다 신중하게 접근해야 한다는 것을 말해준다.

첫째, 증거기반 연구는 효과성 여부만을 따지기 때문에 제도나 프로그램이 실제 작동하는 맥락에 대한 이해가 부족하다. Pearce와 Raman(2014: 395)은 “랜덤통제실험(RCT)이 무엇이 효과가 있는지 일반적 증거를 제시할 때, 정책의 실행과 맥락에 담긴 조건이나 가정은 사라진다”고 비판하였다. 정책이 실행되는 맥락은 무시한 채 하나의 사례에서 입증된 효과가 다른 사례에 그대로 적용될 것이라는 안이한 태도를 지적한 것이다. 어떤 프로그램이 특정 장소와 특정 시점에서 효과가 있다는 증거가 발견되더라도 정책 결정에는 다른 장소와 다른 시점에서도 작동할 것이라는 증거가 필요하다. 무엇이 작동하는지에 대한 증거뿐만 아니라 관련된 여러 사례에도 적용이 가능한 증거가 요구된다(Stoker and Evans, 2016). 이런 지적은 국내 소년사법정책의 증거기반 정책을 수립하기 위해 해외 연구에 전적으로 의존해야 하는 상황에서 문화적 배경이나 사회적 맥락이 다른 해외 연구를 어느 수준까지 차용할 수 있을지에 대한 고민과 연결된다(김병배, 2017). 아울러 증거기반이 추구하는 일반화된 증거가 집단 전체에 적용할 때와 달리 개별 사례에 적용할 때 결과가 다르게 나타날 가능성도 문제로 지적된다(공계순·서인해, 2006). 아무리 외부에서 훌륭히 검증된 증거라도 개별 사례에는 맞지 않을 수 있기 때문

이다.

둘째, 증거기반이 강조하는 위계적인 ‘증거’의 기준이 사회과학 방법론의 다양성을 해치고 획일화를 가져올 수 있다는 우려가 제기된다. 이는 증거기반과 관련하여 가장 비판을 많이 받는 내용으로 최근에는 증거기반 옹호자 중에도 방법론적 다양성을 인정하자고 주장한다(한승훈·안혜신, 2021). 의학에서도 일찍부터 증거에 대한 논란이 일어났다. Sackett과 동료들(1996: 72)은 증거기반의학의 필요성은 인정되지만 증거만으로 치료행위를 정당화할 수 없고, 임상경험에서 비롯된 의사 개인의 지식도 함께 고려해야 한다고 보았다. 증거기반 관점에서 더 확실하고, 더 정확하고, 더 효과적이며, 더 안전한 치료법이 기존 임상경험의 지식을 대체할 수 있지만, 그렇다고 외부적 증거만 강조하는 것은 또 다른 ‘증거기반 독재(becoming tyrannised by evidence)’를 불러올 수 있다. 특히 사회과학에서 연구의 질은 측정절차와 측정도구에 따라 달라질 수 있고, 타당한 측정도구가 무엇인지도 논란이 있을 수 있다. 랜덤통제실험을 최우선 증거로 강조하는 것은 평가기준 중에서 신뢰도와 타당도만 중요하게 여기고 연구목적에 부합하는 적합성이나 연구의 실현가능성은 배제한 기준이라는 비판도 제기된다(공계순·서인해, 2006). Pearce와 Raman(2014: 396)은 랜덤통제실험의 일방적 추구가 더 광범위한 기준을 사용해 가장 적합한 정책을 결정하는 실용적 탐구(practical quest)보다 불확실성을 줄이기 위한 ‘기술적 실행(technical exercise)’에 연구자의 전문성을 제한시킬 수 있다는 사실에 우려를 나타낸 바 있다.

셋째, 증거기반이 객관적인 합리성을 강조하고 있으나 그 자체로 가치중립적이거나 이해관계에서 자유로울 수 없다는 비판이다. Witkin와 Harrison(2001)은 무엇이 증거가 되고 누가 그것을 정하는지, 증거기반을 통해 누가 이득을 보고 손해를 보는지는 ‘힘의 관계’에서 나오는 것이라고 주장하였다(공계순·서인해, 2006에서 재인용). 증거기반 의학이 소개되고 관심이 집중되던 당시 Spence(2014)는 제약회사들이 증거기반을 이용하여 의사의 재량과 판단을 허용하지 않고 “증거가 말하는 대로(사실상 자신들이 원하는 대로)” 과다처방과 과잉진료의 문제를 일으켰다고 비판하였다. 그는 국립보건연구원과 코크란 재단마저 제약회사와 이해관계가 있는 저자를 배제하지 않았다면서 증거기반은 오염되었을 뿐 아니라 결함이 많다고 하였다. Pearce와 Raman(2014: 397)도 랜덤통제실험을 중시하는 증거기반이 제대로 실현되기 위해서는 제도적으로 반드시 ‘투명성

(transparency)’이 보장되어야 한다고 하였다. 투명성과 기준을 잃어버리면 과학은 더 이상 과학이 아니기 때문이다. 증거기반을 추구하는 학자들이 불확실성을 해결하기 위해 고도로 형식화된 양적 방법을 사용할 때, 주어진 정책 문제가 의심스럽고 정치이념에 편승한 것이라 판단이 들면 이미 ‘주어진’ 틀을 단순히 받아들이는 대신 대안적인 정책 프레임에 관여할 필요가 있다고도 하였다. 증거기반은 어떤 면에서 정치와 완전 분리된 것으로 보기 어렵기 때문이다(Pearce and Raman, 2014). 실제로 의도를 가진 정책당국이나 특정집단이 제기한 문제에 집중하고 관련된 증거를 찾는데 매몰되다 보면 정작 그 분야에서 가장 중요한 이슈를 놓칠 수 있다는 점에서(공계순·서인해, 2006), 이러한 비판은 경청할 필요가 있다.<sup>8)</sup> 예를 들어, 보호관찰이나 교정 분야에서 재범율 감소에 효과가 있다는 증거는 모두 합리화될 수 있는 것인가, 재범율 감소가 소년사법제도의 유일한 목표인가, 법원이 명령하는 소년보호 처우는 제대로 작동하고 있는가(김병배, 2017) 등의 문제는 현재 뜨거운 논란이 되고 있는 연령하향 문제만큼이나 중요하지만 정치권이나 언론, 시민은 여기에 별 관심을 보이지 않는다.

이러한 비판에도 불구하고 증거기반은 여전히 매력적이다. 근대 이후 역사의 중요한 분기점마다 타당한 지식과 증거의 축적은 인간의 ‘자유(liberty)’를 증진하는 데 이바지했기 때문이다. 다만, 메타분석이 말하는 ‘forest plot’ 같은 증거는 그 자체로 정책의 결론이 되거나 정책을 결정하지는 않는다. 오히려 증거는 현재까지 알려진 내용 중에서 어떤 개입이 효과적이고 어떤 개입이 덜 효과적인지 정확한 정보를 제공할 뿐이다(Sherman, 2009). 정책 입안자 입장에서도 모든 정책을 정치적으로 결정하기란 쉽지 않다. 정책결정 과정에서 과학적 지식과 증거가 있다면, 정책의 방향을 설정한 후 세부적인 정책추진 과정에서 과학적 지식과 증거에 도움을 받을 수 있다(한승훈·안혜선, 2021).

그러므로 증거기반을 무조건 맹신할 필요도, 무조건 배척할 필요도 없다. 증거기반 형사정책의 필요성을 강조한 Welsh와 Farrington(2001)은 당시까지 알려진 범죄학이론에 기초한 범죄예방 프로그램 중에서 효과가 검증된 것과 그렇지 않은 것을 소개하면서 아무리 중요한 범죄예방 프로그램이라도 국방, 환경, 보건 등 다른 분야에 밀려 후순위가 될 수 있음을 언급하였다. 정책에 반영되거나 채택되는 것과 무관하게 증거에 기반한 연

8) 촉법소년 처벌강화와 연령하향에 대한 논란은 2007년 소년법 개정 전후부터 꾸준히 있었으나, 최근 정부와 정치권이 선거공약, 정책현안, 입법제안을 통해 이슈화하면서 더욱 부추겨진 감이 없지 않다.

구를 꾸준히 진행하고 축적하는 것이 중요하다는 점을 강조한 일화이다. 형사정책 분야는 이미 무엇이 효과가 있고 없는지 알고 있는데도 적절한 증거가 부족한 경우가 너무 많다. 어느날 갑자기 아무도 관심을 두지 않았던 정책에 갑자기 높은 관심이 쏟아질 때, 축적된 증거가 없다면 확실한 증거가 없는데도 옹호자들은 자신의 신념을 뒷받침하기 위해 거짓 증거를 제시하거나 약한 증거를 부풀릴 수 있다(Sherman, 2009: 13). 다른 측면에서 증거기반 연구는 실제 정책으로 반영되지 않더라도 현재 논란이 되고 있는 사안이 무엇인지 명확히 보여주고 논쟁을 불러일으킨다는 점에서 긍정적인 기능을 수행한다(Tonry, 2010: 794). 증거기반이 정책결정에 곧바로 영향을 미칠 것이라는 기대는 순진한 생각일 수 있으나, 사형제도 철폐와 같이 논쟁적 주제이든 방법론 설치와 같이 분명한 주제이든 증거기반은 정책 소통의 장으로 사람들을 모으고 다양한 의견을 교환할 수 있게 한다는 점에서 ‘공공 범죄학(public criminology)’이 지향하는 가치와도 맞닿아 있다(Clear, 2010; Uggen and Inderbitzin, 2010).<sup>9)</sup>

이 논문은 최근 여론의 집중적인 관심을 받은 소년범죄의 자연령화, 흉포화 담론에 따라 형사처분 가능 연령을 낮춰야 한다는 주장에 대한 증거기반 연구의 필요성을 강조하기 위해 미국에서 수행된 책임능력에 관한 연구, 그리고 형사이송제도와 같은 엄벌주의 정책의 효과를 다룬 연구를 체계적으로 검토하였다. 그러나 형벌 제재가 아닌 형벌 이외 제재(소년법의 보호처분)를 다룬 연구, 예를 들면 교화·교육 위주의 보호처분 효과나 형사제재 및 보호처분의 효과를 비교한 연구를 검토하지 못한 한계가 있다. 무엇보다 소년범죄자의 책임능력에 관한 미국의 다양한 사례를 검토했음에도 불구하고 최근 논란의 핵심이 되었던 ‘형사미성년자 연령 하향(14세→13세)’에 관한 직접적이고 분명한 증거는 발견하지 못했다. 이 논문에서 살펴본 ‘책임능력관련 연구’의 상당수는 청소년기가 아동기나 성인기와 무엇이 다른지에 관한 간접적 연구결과일 뿐 국내 논쟁의 핵심이 되고 있는 13세와 14세의 책임능력 차이를 보여주는 연구라고 보기는 어렵다. 비록 연구자들의 한계와 지면의 제약으로 인해 이 논문에서는 논쟁의 직접적인 증거가 될 만한

9) 공공 범죄학과 관련하여 Tonry(2010)는 범죄학사를 거론하면서 최근의 논의가 정말로 새로운 것인지 회의적인 입장을 내비치기도 한다. 그는 1930~40년대 사형 논쟁에 참여했던 Thorstein Sellin, 1950~60년대 시카고 지역 프로젝트에 참여했던 Shaw와 McKay, 1950년대 양형 개혁의 기초자료를 제공하기 위해 비행소년 중단연구를 수행한 Sheldon Glueck 등 초창기 범죄학 역사를 되돌아볼 때 범죄학은 이미 오래전부터 ‘증거기반’과 ‘공공범죄학’을 추구했다고 반박한다.

사례를 충분히 다루지 못하였으나, 추후 연구에서는 엄벌주의 정책의 한계를 보여줄 수 있는 보다 직접적인 연구성과가 소개될 수 있기를 기대한다.

국내에도 소년사법정책과 관련해 증거기반을 추구하는 일부 연구가 수행되고 있으나(김병배, 2017; 박성훈 외, 2020), 질적 측면은 차치하더라도 양적 측면에서도 증거기반 연구가 충분하지 않은 것이 사실이다. 국내에서 증거기반 형사정책을 실현하기 위한 대안으로는 각 분야마다 인과적 실험설계 형태를 갖춘 연구의 꾸준한 생산과 축적, 기존 연구의 수집 및 체계적인 분석을 위한 증거기반 전담 연구조직의 마련, 범죄자에 대한 대규모 코호트 및 패널자료의 구축과 공개, 대통령실·법무부·경찰·검찰·법원 등 정부기관의 정치적 이니셔티브 등이 언급되고 있다(박철현, 2014; 서홍란, 2014; 박미량, 2014; 김병배, 2017). 기존 연구에 대한 메타분석을 정기적으로 실시하고 그 결과를 확산하는 정책연결형 연구체계 구축에 있어 한국형사·법무정책연구원에 전담부서를 설치하는 방안(박철현, 2014; 김병배, 2017)도 고려할 만한 대안으로 여겨진다.

끝으로 형사미성년자 연령을 둘러싼 논쟁을 정리하면서 국내에 증거기반 형사정책 개념이 일찍이 소개되고 그 필요성이 강조된 후에도 여전히 관련된 증거기반 연구가 부족하다는 사실에 안타까움을 느꼈다. 그러나 다른 한편으로 형사정책의 중요한 한 축을 담당하는 범죄학의 학문 정체성이나 다양한 방법론 자체가 이미 증거기반과 친화성이 강하다면, 우리 학계의 학문적 역량과 관계 기관들의 높은 관심을 고려할 때 머지않아 국내에서도 여러 분야의 수많은 증거기반 연구가 축적되리라 기대한다. 이를 위해서는 향후 증거기반 형사정책 분야에서 무엇을 증거로 볼 것인지, 형사정책 분야의 증거기반이 무엇을 의미하는지, 어떤 형사정책을 증거기반으로 연구해야 하는지, 증거기반 형사정책은 무엇을 추구해야 하는지에 대한 활발한 토론이 계속해서 이어져야 할 것이다.

## 참고문헌

- 공계순, 서인해. (2006). 증거기반 사회복지실천에 대한 이해와 한국에서의 적용가능성에 관한 연구. *사회복지연구*, 31, 77-102.
- 김두상, 박상식. (2016). 형사처벌 대상자의 연령인하와 치유사법에 관한 고찰. *법학연구*, 24(2), 1-20.
- 김수영. (2001). 근거중심의학. *연세의학교육*, 3(1), 1-19.
- 김봉수. (2013). 소년범죄의 최근 동향과 대책. *법학논총*, 33, 173-197.
- 김병배. (2017). 소년 보호관찰 분야에서의 증거기반 정책 채택과 향후 과제. *보호관찰*, 17(1), 41-81.
- 김연지. (2018). 소년의 형사책임능력과 치료적 처우 방안에 관한 고찰. *비교형사법연구*, 20(2), 171-195.
- 남선모. (2019). 형사책임연령 하향조정 인식에 따른 실질적 개선에 관한 연구. *인문사회과학연구*, 27, 32-52.
- 박미랑. (2014). 미국의 소년범 형사이송제도의범죄 억제력에 관한 고찰. *형사법의 신통향*, (45), 106-132.
- 박상식. (2013). 소년범죄와 학교폭력 예방을 위한 회복적 사법의 도입에 관한 연구. *법학연구*, 21(2), 211-241.
- 박상식. (2017). 소년범의 연령하향과 처벌강화에 관한 소고. *법학연구*, 25(4), 99-126.
- 박선영. (2019). 소년사범에 대한 형사처분의 효과성 진단-미국의 평가연구 분석. *소년보호연구*, 32(1), 105-130.
- 박성훈, 전영실, 정진경. (2020). 소년원생을 위한 사회정착프로그램의 단계별 시범운영 및 효과성 분석. *소년보호연구*, 33(1), 87-112.
- 박수희, 문준섭. (2018). 형사미성년자 연령의 하향 개정논의에 대한 비판적 검토. *한국경찰학회보*, 20(6), 143-172.
- 박지혜, 이수정. (2022). 형사미성년자 기준연령 하향에 대한 고찰. *교정담론*, 16(1), 221-251.
- 박찬걸. (2013). 형법상 형사미성년자 연령 설정과 소년법상 소년보호처분제도와의 관

- 계. 소년보호연구, 22, 191-222.
- 박찬걸. (2020). 제20대 국회에 제출된 소년법 개정법률안에 대한 검토: 제재강화에 대한 비판을 중심으로. 형사정책, 32(2), 149-178.
- 박철현. (2014). 증거에 기반한 형사정책의 발전과 국내 적용방향. 형사정책연구, 25(1), 123-157.
- 백영민, 박인서. (2020). R기반 성향점수분석: 루빈 인과모형 기반 인과추론. 한나래.
- 손현중, 성진기. (2023). 촉법소년 연령 정책에 관한 정책분석과 시사점 연구: Kingdon의 정책 흐름 모형 중심으로. 범죄수사학연구, 9(1), 111-132.
- 서홍관. (2014). 보호관찰 지도감독에서의 증거기반실천의 적용. 보호관찰, 14(2), 7-51.
- 오철호. (2015). 정책결정, 증거 그리고 활용: 연구경향과 제언. 한국정책학회보, 24(1), 53-76.
- 원혜욱. (2022). 촉법소년 상한 (형사미성년자) 연령 하향에 대한 비판적 고찰. 소년보호연구, 35(2), 167-192.
- 이덕인. (2012). 형사책임연령 하향에 대한 비판적 고찰: 형사미성년과 촉법소년을 중심으로. 형사정책연구, 5-40.
- 이용식. (2017). 소년 위법행위자의 연령에 관한 몇 가지 소고. 소년보호연구, 30(3), 97-121.
- 이인영. (2015). 청소년 뇌과학 연구가 미국의 소년사법에 미친영향에 관한 고찰. 홍익법학, 16(3), 433-455.
- 임민경, 이지혜, 이한나, 김태동, 최기홍. (2013). 근거기반실천과 심리치료. 한국심리학회지: 일반, 32(1), 251-270.
- 정신교, 차시환. (2014). 소년의 형사책임연령에 대한 재검토. 법학연구, 22(4), 203-228.
- 점승현. (2022). 촉법소년의 연령 하향. 법이론실무연구, 10(3), 345-367.
- 하태훈. (2022). 증거기반 형사정책과 형사입법. 형사정책, 34(3), 7-34.
- 한국리서치. (2023). 여론 속의 여론: 촉법소년 범죄 해결 여론. 한국리서치 주간 리포트, 215-2.
- 한승훈, 안혜선. (2021). 증거기반 정책의 쟁점과 한국적 맥락에서의 적용 가능성. 한국정책학회보, 30(1), 289-314.

- Baerger, D. R., Griffin, E. F., Lyons, J. S., & Simmons, R. (2003). Competency to stand trial in preadjudicated and petitioned juvenile defendants. *Journal of the American Academy of Psychiatry and the Law Online*, 31(3), 314-320.
- Cascio, C. J., Gerig, G., & Piven, J. (2007). Diffusion tensor imaging: application to the study of the developing brain. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(2), 213-223.
- Casey, B. J., & Jones, R. M. (2010). Neurobiology of the adolescent brain and behavior: implications for substance use disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(12), 1189-1201.
- Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, 1124, 111-126.
- Cauffman, E., & Steinberg, L. (2000). (Im) maturity of judgment in adolescence: Why adolescents may be less culpable than adults. *Behavioral sciences & the law*, 18(6), 741-760.
- Clear, T. R. (2010). Editorial Introduction to Public Criminologies. *Criminology & Pub. Pol'y*, 9, 721.
- Cohen Gilbert, J. E., & Thomas, K. M. (2013). Inhibitory control during emotional distraction across adolescence and early adulthood. *Child development*, 84(6), 1954-1966.
- Cooper, D. K. (1997). Juveniles' understanding of trial related information: Are they competent defendants?. *Behavioral Sciences & the Law*, 15(2), 167-180.
- Cullen, F. T., & Jonson, C. L. (2016). *Correctional theory: Context and consequences*. Sage Publications.
- Dreyfuss, M., Caudle, K., Drysdale, A. T., Johnston, N. E., Cohen, A. O., Somerville, L. H., Galvan, A., Tottenham, N., Hare, T. A., & Casey, B. J. (2014). Teens impulsively react rather than retreat from threat.

- Developmental neuroscience, 36(3-4), 220-227.
- Epstein, R. (2007). The myth of the teen brain. *Scientific American Mind*, 18(2), 56-63.
- Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., Blair, J., & Pine, D. S. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage*, 25(4), 1279-1291.
- Fagan, J., Kupchik, A., & Liberman, A. (2004). Be careful what you wish for: Legal sanctions and public safety among adolescent offenders in juvenile and criminal court. Columbia Law School, Pub. Law Research Paper, (03-61).
- Fowler, E., & Kurlychek, M. C. (2018). Drawing the line: Empirical recidivism results from a natural experiment raising the age of criminal responsibility. *Youth violence and juvenile justice*, 16(3), 263-278.
- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of neuroscience*, 26(25), 6885-6892.
- Galvan, A., Hare, T., Voss, H., Glover, G., & Casey, B. J. (2007). Risk taking and the adolescent brain: Who is at risk?. *Developmental science*, 10(2), F8-F14.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.
- Levitt, S. D. (1998). Juvenile crime and punishment. *Journal of political Economy*, 106(6), 1156-1185.
- Liston, C., Watts, R., Tottenham, N., Davidson, M. C., Niogi, S., Ulug, A. M., & Casey, B. J. (2006). Frontostriatal microstructure modulates efficient recruitment of cognitive control. *Cerebral cortex*, 16(4), 553-560.

- Mackenzie, D. Layton. (2000). Evidence-based corrections: Identifying what works. *Crime and Delinquency* 46(4), 457-471.
- Millenson, M. (1997). *Demanding Medical Excellence: Doctors and Accountability in the Information Age*. Chicago, IL: University of Chicago Press.
- Pearce, W., & Raman, S. (2014). The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance. *Policy sciences*, 47, 387-402.
- Sackett, D. L., & Rosenberg, W. M. C. (1995). The need for evidence-based medicine. *Journal of Royal Society of Medicine*, 88, 620-624.
- Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *Bmj*, 312(7023), 71-72.
- Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence-based medicine. *Journal of Public Health*, 17(3), 330-334.
- Shavit, Y., & Rattner, A. (1988). Age, crime, and the early life course. *American Journal of Sociology*, 93(6), 1457-1470.
- Sherman, L. W. (1998). *Evidence-based policing* (p. 15). Washington, DC: Police Foundation.
- Sherman, L. W. (2009). Evidence and liberty: The promise of experimental criminology. *Criminology & Criminal Justice*, 9(1), 5-28.
- Somerville, L. H., Hare, T., & Casey, B. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *Journal of cognitive neuroscience*, 23(9), 2123-2134.
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature neuroscience*, 6(3), 309-315.
- Spence, Des. (2014). Evidence based medicine is brokes. *British Medical Journal*, 348.

- Stoker, G., & Evans, M. (2016). Evidence-based policy making and social science. In *Evidence-Based Policy Making in the Social Sciences* (pp. 15-28). Policy Press.
- Tonry, M. (2010). Public criminology and evidence-based policy. *Criminology & Pub. Pol'y*, 9, 783.
- Uggen, C., & Inderbitzin, M. (2010). Public criminologies. *Criminology & Public Policy*, 9(4), 725-749.
- Venturelli, A. (2021). Young Adults and Criminal Culpability. *U. Pa. J. Const. L.*, 23, 1142.
- Walsh, A., & Beaver, K. M. (2009). *Biosocial criminology* (pp.79-101). Springer New York.
- Welsh, B. C., & Farrington, D. P. (2001). Toward an evidence-based approach to preventing crime. *The Annals of the American Academy of Political and Social Science*, 578(1), 158-173.
- Wilson, J. Q., & Herrnstein, R. (1985). *Crime and human behavior*. New York: Si Mon Schuster.
- Winner, L., Lanza-Kaduce, L., Bishop, D. M., & Frazier, C. E. (1997). The transfer of juveniles to criminal court: Reexamining recidivism over the long term. *Crime & Delinquency*, 43(4), 548-563.
- Witkin, S. L., & Harrison, W. D. (2001). Whose evidence and for what purpose?. *Social Work*, 46(4), 293-296.
- Zane, S. N., Welsh, B. C., & Mears, D. P. (2016). Juvenile transfer and the specific deterrence hypothesis: Systematic review and meta analysis. *Criminology & Public Policy*, 15(3), 901-925.

## The Review and Prospect for Evidence-Based Researches on Punitive Juvenile Justice Policy

Park, Seong-Hoon\* · Park, Sin-ui\*\*

This research focused on the issue of lowering the minimum age of criminal responsibility, which has been a hot political issue in South Korea. Reviewing domestic and foreign literature, we comprehensively examined the main grounds of each side. With this process, we found that debates have been too normative or have intensified the confusion through some evidence. To get through it, we covered the concept and the meaning of ‘evidence-based.’ Also, we reviewed American evidence-based research on the age of criminal responsibility from two aspects. First, we examined biological evidence such as brain maturation and neuroscience to see whether the juvenile offender is not indistinctive from adult offenders in doli capax. Second, we looked into empirical studies on juvenile transfer to determine whether reinforcing punishment for juvenile offenders brings about better outcomes like reducing recidivism rates. We carefully concluded that the negative impacts of lowering the minimum age of criminal responsibility would outweigh the positive impacts. Therefore, efforts to accumulate quality evidence are required with a prudent recognition that an evidence-based approach is not a panacea.

❖ Key words: Juvenile Delinquency, Criminal Minors, Doli Capax, Evidence-based Policy, Randomized Controlled Trial (RCT)

투고일 : 7월 26일 / 심사일 : 9월 22일 / 게재확정일 : 9월 30일

---

\* Senior Research Fellow, Ph.D. Korean Institute of Criminology and Justice

\*\* Doctoral Student, Department of Criminal Justice, University of Cincinnati