creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

사회학 석사 학위논문

정신질환 연구의 유전화는 누가 주도하는가

- 정부와 기업을 중심으로 -

2022 년 8 월

서울대학교 대학원

사회학과

최 중 식

정신질환 연구의 유전화는 누가 주도하는가

- 정부와 기업을 중심으로 -

지도 교수 장 덕 진

이 논문을 사회학 석사 학위논문으로 제출함
2022 년 8 월

서울대학교 대학원
사회학과
최 중 식

최중식의 사회학 석사 학위논문을 인준함
2022 년 8 월

위 원 장 _____ 손윤규 _____ (인)

부위원장 _____ 조원광 _____ (인)

위 원 _____ 장덕진 _____ (인)

국문 초록

본 연구는 정신질환(Mental disorder)의 유전화(Genetification)를 유전학(Genetics)과 정신의학(Psychiatry) 분야가 접목하여 정신질환을 유전자와 그들 간의 상호작용의 관점에서 이해하고 설명하려는 현상으로 정의한다. 어떻게 유전학 연구는 정신질환의 진단, 예측, 치료에 대한 포괄적인 의학적 근거를 축적할 수 있었을까? 유전학 연구에 필요한 다양한 인적, 물적, 기술적 자원을 동원하여 유전자를 둘러싼 과학적 지식을 생산하는 조직적 행위자는 누구일까?

해당 질문에 답하기 위해, 본 연구는 유전화 현상을 주도할 수 있는 행위자로서 정부와 기업(제약 및 바이오 기술 회사)이 정신질환의 의료 지식에 미치는 영향력을 분석하였다. 본 연구는 펍메드(PubMed)에서 정신질환을 다루는 논문을 수집하고, 구조적 토픽모델링(Structural topic model)을 활용하여 논문의 출간 시기, 정부의 펀드 여부, 저자의 기업 소속 여부가 텍스트의 토픽 발현 확률에 어떻게 영향을 주는지 분석하였다.

분석 결과는 다음과 같다. 첫째, 정부는 뇌와 유전자로 대표되는 정신질환의 생물학적 원인을 파악하는 것에 집중한다. 둘째, 정부는 환자의 신체적 증상을 바탕으로 환자 집단을 구성하고, 그들이 사회적 관계 안에서 어떤 고통을 마주하는지 주목한다. 셋째, 기업은 약물의 효과 및 안전성을 검증하는 것에 초점을 둔다. 넷째, 출간 시기 변수를 통제했을 때, 유전학 연구는 정부에 의해 주도되지만, 출간 시기와 기업 소속 변수의 상호작용을 관찰했을 때, 유전학 연구는 2010년 이후 기업 소속의 논문에서 그 비중이 증가했다.

논문의 분석 결과는 두 가지 함의를 갖는다. 첫째, 정부와 기업은 정신질환에 대한 의료 지식을 둘러싸고 일종의 역할 분담을 수행한다. 정부가 정신질환의 원인과 증상에 주목한다면, 기업은 약물 치료의 안전성과 효과에 집중한다. 둘째, 유전학은 정부와 기업의 이해관계가 결합할 수 있는 분야이다. 정부는 유전적인 위험 인자에 근거하여 정신질환을 조기에 효과적으로 관리할 수 있다. 기업은 부작용을 줄이고 유전자의 기전을 조기에 억제할 수 있는 맞춤형 약물을 개발하여 의료 시장을 확대할 수 있다.

주요어 : 정신질환, 유전화, 구조적 토픽 모델링, 정부, 기업
학 번 : 2018-28122

목 차

제 1 장 들어가며 1

제 2 장 이론적 배경과 선행연구 검토 3
제 1 절 정신질환 연구의 유전화 3
제 2 절 생의료화 9

제 3 장 자료 및 연구 방법 14
제 1 절 자료와 변수 14
제 2 절 구조적 토픽모델링 18
제 3 절 전처리 및 통계 도구 21

제 4 장 분석 결과 22

제 5 장 마치며 34

제 6 장 부록 37

참고문헌 44

Abstract 47

그림 목차

[그림 1]전체 논문의 수 추이 16
[그림 2]구조적 토픽모델링의 생성 과정 19
[그림 3]진단통계치 23
[그림 4]1, 18번 토픽의 출간 시기에 따른 발현 확률의 변화 . 25
[그림 5]7, 9, 11, 16번 토픽의 출간 시기에 따른 발현 확률의 변화 27
[그림 6]정부/비정부 변수와 토픽 발현 확률의 관계 29
[그림 7]기업/비기업 변수와 토픽 발현 확률의 관계 30
[그림 8]1번 토픽에서 기업/비기업 변수와 출간 시기의 상호작용 효과 34

표 목차

[표 1]변수에 따른 데이터 분류 행렬 17

제 1 장 들어가며

질병은 어떻게 유형화되느냐에 따라서 새로운 사회적 담론을 만들어낸다 (조병희, 2006). 유전학 연구들은 다양한 질병들의 유전적 원인에 대해 밝히면서 질병을 어떻게 정의하고 치료할 것인가에 대해 새로운 의학적 시선을 만들고 있다. 비만을 예로 들어보자. 지금까지 비만은 일련의 증상만 존재하는 증후군으로서, 생활습관이 비만의 주요 원인으로 지적되었다. 하지만, 비만 유전자가 발견된다면 얘기가 달라진다. 기존에는 생활습관을 바꾸는 것이 중요했다면, 이제 비만은 유전자 치료를 통해 극복해야 하는 질병이 된다. 즉, 외현적 증상이 아닌 유전자가 질병을 정의하고, 분류하고, 치료하는 것에 있어 중요한 요소가 된다.

본 연구는 유전학(Genetics)과 정신의학(Psychiatry) 분야가 결합하여 정신질환(Mental disorders)을 유전자와 그들 간의 상호작용의 관점에서 이해하고 설명하려는 현상으로서 정신질환의 유전화(Genetification)에 주목한다. 의학적 지식은 인구의 정상과 비정상을 나누는 기술을 제공하고 비정상을 치료하여 사회의 정상적 표준을 유지할 수 있도록 한다 (조병희, 2006). 정신의학은 소위 “광인” 입구 집단을 분류하고, 관리하고, 치료하기 위해 “무엇이 인간의 정신을 병들게 하는가”라는 질문에 답해야 했다. 19세기 말 과학자들이 뇌의 어떤 화학적, 유전적 변화가 정신질환을 발생시키는지 연구하기 시작하면서 정신질환은 의료적 치료의 대상이 되었다. 하지만, 정신질환은 세대를 통해 후세로 전달되면서 그 가계에 속한 사람들을 퇴화시킨다는 “퇴행이론”이 유럽 사회를 휩쓸게 되면서, 유전학 연구는 특정 인종, 범죄자, 일탈자 등의 인구 집단을 사회에서 억압하고 퇴출하는 과학적 근거를 제공하기도 하였다 (Shorter, 2020).

그 후, 1970년대에 분자 생물학이 발전함에 따라 정신질환이

선천적으로 결정될 수 있다는 주장에 대한 보다 과학적 근거들이 발견되기 시작했다. 최근 과학자들은 정신질환을 일으키는 유전적인 위험 인자를 발견하고, 이를 사전에 억제할 수 있는 약물적 치료 방식에 대해 연구하고 있다. 정신질환의 유전화 현상은 마음이라는 추상적인 신체의 기전조차 선천적으로 결정될 수 있고, 아직 증상이 나타나지 않더라도 예방적 치료의 대상이 될 수 있다는 새로운 의학적 시선을 만들어내고 있다. 예컨대 유전학의 성과에 근거하여 자신의 정신질환을 관리하려는 소비자층이 형성되고, 유전자 연구에 대한 다양한 법적, 윤리적 논쟁이 발생하고 있는 현실은 새로운 의학적 시선의 사회적 영향력을 보여준다.

본 연구는 어떤 조직적 행위자가 정신질환의 유전화를 추동하고 있는지 분석하고자 한다. 의료의 본질적인 목적은 건강을 증진하고 질병을 관리하는 것이지만, 의료는 동시에 의료 산업이 되고 공급자의 이해관계가 작용하게 된다 (조병희, 2006). 정신질환 연구를 구성하는 세부 분야 중에서 왜 유전학 연구에 많은 학문적 자원이 투입되는지, 그에 따라 어떻게 유전학은 진단, 예측, 치료에 대한 포괄적인 의학적 근거를 축적할 수 있었는지 답하기 위해서는 지식 바깥의 요소에 주목할 필요가 있다. 본 연구는 정신질환의 유전화 현상에 정부와 기업(제약회사)의 이해관계가 반영되어 있다고 본다. 정부는 질병의 발병에 기여하는 유전적 위험 인자를 파악하고 이를 사전에 통제하여 질병을 조기에 관리하고 예방할 수 있다. 기업은 정신질환을 유발하는 유전자의 신체적 기전에 직접적으로 개입할 수 있는 약물을 개발하여 더 많은 의료 소비자들을 확보할 수 있다.

본 연구는 펍메드(PubMed)에서 정신질환을 다루는 논문을 수집하고 구조적 토픽모델링(Structural topic model)을 활용하여 정부와 기업이 정신질환의 의료지식에 미치는 영향을 분석하였다. 구체적으로, 논문의 출간 시기, 정부의 펀드 여부, 저자의 기업 소속 여부가 텍스트의 토픽 발현 확률에 어떻게 영향을 주는지 분석하였다. 그 결과, 다음과 같은 분석 결과를 도출하였다. 첫째, 정부는 뇌와 유전자로

대표되는 정신질환의 생물학적 원인을 파악하는 것에 집중한다. 둘째, 정부는 환자가 겪는 신체적 증상을 바탕으로 환자 집단을 구성하고, 그들이 사회적 관계 안에서 어떤 고통을 마주하는지 주목한다. 셋째, 기업은 환자가 겪는 신체적, 정신적 증상이 아닌, 약물의 효과 및 안전성을 검증하는 것에 초점을 둔다. 넷째, 논문의 출간 시기와 기업 소속 변수의 상호작용을 분석했을 때, 유전학을 다루는 논문들은 2010년 이후 기업 소속의 논문에서 그 비중이 증가했다. 반면, 기업과 연계되지 않은 논문에서는 그 비중이 거의 같은 수준으로 유지되었다. 이는 기업의 이해관계가 정신질환의 세포 병리에 근본적으로 개입할 수 있는 치료방법 및 약물을 개발하는 쪽으로 발전하고 있다는 것을 보여준다.

이어지는 장에서는, 정신질환이 어떻게 의료적 차원의 문제로 연구되기 시작했고 정신의학과 유전학의 학제 간 연구가 정신질환 연구의 지형을 어떻게 바꾸고 있는지 논의할 것이다.

제 2 장 이론적 배경과 선행연구 검토

제 1 절 의료화와 정신질환 연구의 유전화

본 연구는 현재 광범위하게 받아들여지고 있는, 콘래드가 제시한

의료화(medicalization) 개념의 정의에서부터 출발한다. 그는 의료화를 “특정 문제가 질병이나 질환으로 정의되고 의학적 개입을 통해 치료되는 것, 혹은 비의료적 문제가 질병이나 질환과 같은 의학적 문제로 치료되는 과정”(Conrad, 2005)으로 정의한다. 정신질환이 의료적 문제로 인식되기 전까지 정신질환 환자들은 가정이나 수용소와 같은 시설에 감금해야 하는 대상이었다. 하지만, 수용소 바깥의 연구소와 대학에서 과학자들이 뇌와 중추신경계의 어떤 유전적, 화학적 변화가 인간의 정신을 병들게 하는지 질문하기 시작하면서, 정신질환은 의학적 문제로 간주되기 시작했다. 이처럼, 인간의 정신질환을 신경과 뇌의 질환으로 인식하는 의학적 관점을 생물학적 정신의학이라고 한다. 19세기의 생물학적 정신의학은 주로 뇌를 해부해서 정신질환의 물리적 원인을 찾는 것에 집중했다 (Shorter, 2020).

쇼터(2020)는 19세기의 생물학적 정신의학이 쇠퇴한 원인으로 두 가지를 꼽는다. 첫째, 19세기 정신과 의사들은 정신질환의 생물학적 이해를 높이기 위해 정신질환의 유전적 요인에 주목하여 정신질환이 어떻게 세대를 통해 후대로 전달되는지 연구했다. 그들의 생각은 퇴행이론으로 발전하였는데, 퇴행이론에 따르면 정신질환이 세대를 통해 후세로 전달되면서 그 가계에 속한 사람들은 점진적으로 열등한 존재로 퇴화된다. 퇴행이론의 지지자들은 환자 가족이 속한 인구집단 전체가 퇴행에 빠질 수 있기 때문에, 사회의 진보를 위해 이들을 사회에서 축출해야 한다는 사회적 의제를 형성하였다. 20세기 초 퇴행이론은 점차 유럽 문화에서 중요한 부분을 차지하게 되었고, 우생학자와 사회위생운동가들이 퇴행이론을 차용하게 되면서, 퇴행이론은 정신지체, 동성애, 유태인과 같은 “퇴행자” 집단을 단종해야 한다는 국가폭력의 논리와 결합하게 되었다. 결국, 나치즘이 유전론을 악용했던 2차 대전 이후 유전학에 관한 학문적 토론은 일종의 금기 사항이 되었다.

둘째, 뇌 부검을 통해 정신질환의 물리적 병인을 찾아내려는 시도는 현미경에 의존했던 당시의 의학 수준으로 봤을 때 많은 과학적 진척을 내기 어려웠다. 에밀 크레플린은 생물학적 정신의학의 관점과

거리를 두면서 질병의 증상을 상세히 기술하고, 많은 사례를 모아 질병을 체계적으로 분류하는 작업에 집중했다. 그는 정신질환의 유전이나 뇌 생물학 등의 연구로부터 벗어나 시간에 따라 질병이 스스로 드러내 주는 증상에 집중할 것을 요구했다. 즉, 질병의 물리적 원인이 아니라 질병의 진행 과정과 결말, 치료 효과를 의미하는 예후(prognosis)가 중요한 개념으로 등장한 것이다. 크레플린은 질병의 진행 과정에 대한 최대한 많은 사례를 확보해야만 환자들에게 증세가 어떻게 진행될 것인지, 치료를 통해 어떤 효과를 기대할 수 있는지 예측할 수 있다고 봤다.

크레펠린의 의학적 성과는 20세기 정신의학에서 강력한 영향력을 발휘하였다. 그의 관점은 후에 미국 정신의학회의 정신질환 진단 및 통계 편람(Diagnostic and Statistical Manual of Mental Disorders, DSM)의 질병 분류 근거가 되었다. DSM은 어떤 증상들이 얼마나 많이 존재하고, 얼마나 오랜 기간 지속되는지 등에 대한 임상적 결과들을 기반으로 각종 정신질환을 정의하고 분류할 수 있는 기준을 제시한다. 각종 정신질환이 대중적으로 많은 관심을 받게 되면서 DSM은 대중적으로 널리 배포되었지만, 그 모호성 때문에 많은 비판을 받았다 (Frances, 2014). 이러한 비판은 결국 한 질병에 특이적인 증상을 상정하는 것은 매우 어렵다는 정신질환의 근본적인 특성에서 비롯된다. 예를 들어, DSM-5의 진단 기준에 따르면, 주요 우울증은 적어도 2주 이상 지속되는 '우울한 기분' 혹은 '감소된 흥미나 쾌감'이 있으면서 9개의 증상 중에서 5개 이상의 증상이 있는 경우에 진단이 이뤄진다. 경우의 수를 따졌을 때, 이러한 범주를 만족시키는 증상 조합은 227가지에 이른다 (양병환, 2017).

20세기 나치의 잘못된 우생학에 의해 단절되었던 유전학 연구는 현대 분자 생물학의 발전에 힘입어 생물학적 정신의학의 주요 분야로 재등장하였다 (Shorter, 2020). 초기의 유전학 연구는 가족 연구에 집중되었다. 예를 들어, 연구자들은 태어난 직후 완전히 다른 가정에서 자란 쌍둥이들이 정신질환을 공유한다는 사실을 확인함으로써 유전자가

정신질환의 주요 원인 중 하나라는 것을 밝혔다. 그 후 인간의 게놈 구조가 2000년대 초반에 밝혀지고 분자 단위까지 분석할 수 있는 기술이 발달함에 따라서, 인간의 유전적 변이를 보다 효과적으로 탐구할 수 있는 길이 열렸다. 정신질환과 유의성을 갖는 단일한 유전 구조를 찾으려는 시도는 성공적이지 못했지만, 최근의 연구들은 복수의 유전적 변이들과 개인의 생리적 기질, 그리고 환경적 요소들이 상호작용하여 정신질환이 발병한다는 연구결과를 밝히고 있다.

근거기반의료(evidence-based medicine)는 현재까지 알려진 가장 과학적인 연구의 결과를 의학 분야에 적용하여 개별 환자에 대한 의료적 결정을 내려야 한다는 원칙이다. 유전학과 정신의학의 학제 간 연구가 내고 있는 의학적 근거들은 진단, 예측, 치료의 관점에서 새로운 의료 행동을 주도하고 있다. 첫째, 정신의학과 유전학의 만남은 "질병의 원인을 보지 않고 증상에만 의존하는 부드러운(soft) 진단에서 생물학적 원인을 탐구하는 단단한(hard) 진단" (Steven Edward Hyman, 2010)으로의 변화하고 있다는 것을 의미한다. 최근의 연구 결과들은 환자들의 유전 정보와 DSM의 임상적 분류가 높은 상관성을 보이고 있으며, 나아가 유전적 분류가 DSM의 모호성을 극복할 수 있는 대안이 될 수 있다는 것을 보여주고 있다. 예컨대, 유전학 연구는 DSM 체계 내에서는 별개로 분류된 조현병(Schizophrenia)과 양극성장애(Bipolar disorder)가 실제로는 높은 유전적 상관성을 가지고 있는 것을 밝히기도 했다 (Lichtenstein et al., 2009). 둘째, 크레플린 학파가 질병의 경과에 대한 경험적 자료들을 통해 질병의 결과를 예측하고자 했다면, 유전학 연구들은 유전적 변이들이 정신질환의 발병에 어느 만큼의 기여를 하는지 수치적으로 규명하고자 한다. 다시 말해, 질병이 발병하고 그 후의 경과에 대해 말하는 사후 예측이 아니라, 질병이 발병하기 이전에 그 경과가 어떻게 진행될지 예측할 수 있다는 것이다. 셋째, 질병의 발병을 촉진하는 위험 유전자(vulnerability genes)가 언제 활성화되는지 알 수 있다면, 해당 유전자가 활성화되기 전에 그 작용을 사전에 억제할 수 있는 약물을 쓸 수 있다 (Steven E Hyman, 2000; Pezawas et al.,

2005). 최근의 약물유전학(Pharmacogenetics) 연구들은 개인의 유전적 특성에 따라 약물이 신체에 작용하는 효과가 달라진다는 것을 밝히고 있다. 그에 따라, 개인의 유전적 특성에 부합하는 개별화된(personalized) 약물 처방이 가능해진다 (Chang, Weitzel, & Schmidt, 2015; 심창구 & Kyōto Daigaku. Daigakuin, 2012).

지식은 "주목해야 할 대상을 제시하고 그 대상을 인지하는 태도와 방식을 결정함으로써 행동을 유도한다"(조원광, 2017). 정신질환의 유전화 현상은 아직 외현적인 증상이 보이지 않더라도 유전적 위험 인자에 근거하여 잠재적인 환자를 구성하고, 개별화된 치료 방식을 통해 정신질환을 관리해야 한다는 새로운 의학적 시선을 만들어내고 있다. 이러한 새로운 의학적 시선은 정신질환을 둘러싼 행동과 태도를 바꾸고 있다. 유전학의 과학적 발견에 근거하여 진단, 예측, 치료 방식이 발전하고 있으며, 보다 과학적인 근거를 통해 자신의 마음을 관리하려는 소비층이 늘어나고 있다. 유전자 연구가 바꿀 사회의 모습에 대해 다양한 법적, 윤리적 논쟁이 발생하고 있는 현실은 유전학 연구가 많은 사회적 관심을 받고 있다는 것을 보여준다.

하지만, 유전학의 발전이 증상 중심의 진단 기법을 완전히 대체할 수준까지 발견한 것은 아니다. 정신질환의 발병에 기여한다고 알려진 위험 유전자는 수 천개에 이른다. 유전자의 생리적 기전은 다른 유전자와의 상호작용에 의존하고, 상호작용할 수 있는 경우의 수는 무수히 많기 때문에 유전 구조와 정신질환의 유의성을 정확히 밝히는 것은 매우 어려운 일이다. 실제로 한 논문은 많은 표본 집단을 확보했을 때 조현병을 유발한다고 알려진 유전자들이 실제로는 해당 질병과 상관관계가 없다는 것을 밝혔다 (Sanders et al., 2008). 정신질환과 밀접한 관계가 있다고 밝혀진 유전자조차 가족, 학교, 친교 관계 등 개인이 처한 사회적 환경과 상호작용하면서 정신질환의 발병에 기여한다 (Guo, Roettger, & Cai, 2008). 무엇보다, 정신질환은 '마음'의 질병이다. 마음은 인간의 자율적인 주관적 의지가 개입되는 한 "최종적으로 결정된 상태가 아니라 무엇으로 되어가는, 생각하는, 발전해 나가는 것"(양병환,

2017)일 수밖에 없다. 19세기에 생물학적 정신의학이 쇠퇴했던 이유도 정신질환의 병인을 찾느라 정신질환이 마음의 문제라는 점을 간과했기 때문이다 (Shorter, 2020).

그렇다면, 왜 우리는 정신질환의 유전화 현상에 주목해야 할까? 쇼터(2020)는 정신질환이 의학적으로 설명되어 의료화 과정을 거치면서 역설적으로 다양한 의학적 전문 분야가 정신의학의 영역에서 벗어나게 되었다고 지적한다. 정신질환에 대한 약물적 치료의 효과가 발견된 이후, 약물학의 발전은 거의 전적으로 제약회사에 의해 이뤄졌다. 의사-환자의 관계는 정신치료 훈련을 받은 심리상담사, 심리학자, 지역활동가들에게 위협받고 있다. 더불어, 뇌-생물학 분야는 뇌 영상 촬영 결과와 뇌 기저핵 병소를 더 잘 판단할 수 있는 신경과 의사에게 점유될 가능성이 크다. 이는 정신질환에 대한 의료 행위의 기반이 되는 과학적 연구들이 통합적인 의학적 근거를 제공해주고 있지 못한다는 것을 의미한다 (Shorter, 2020). 따라서, 정신의학뿐만 아니라 뇌-생물학, 약물학과 결합하여 정신질환의 진단, 예측, 치료에 대한 포괄적인 의학적 근거를 제공하려 시도하는 유전학의 발전은 주목할 만 하다.

그 불확실성에도 불구하고 어떻게 유전학 연구는 정신질환의 진단, 예측, 치료에 대한 포괄적인 의학적 근거를 축적할 수 있었을까? 의료의 본질적인 목적은 건강을 증진하고 질병을 관리하는 것이지만, 의료는 동시에 의료 산업이 되고 공급자의 이해관계가 작용하게 된다 (조병희, 2006). 정신질환 분야 내에서 유전학 연구는 분자생물학의 기술적 발전, 다양한 유전적 변이를 관찰할 수 있는 충분한 표본 집단, 이들을 실험할 수 있는 충분한 인력과 시설이 요구되는 자본 집약적 특성을 갖고 있다. 이러한 연구는 연구자 개인의 학문적 호기심에서 비롯된 것이라 보기는 어렵다. 정신질환의 유전화를 설명하기 위해서는 다양한 물적, 인적 자원을 확보하고 이를 적재적소에 배치하고 관리할 수 있는 조직적 요소들에 주목할 수밖에 없다. 다음 장에서는 구성주의적 관점에서 지식의 발전에 대해 논하고, 정신질환의 유전화 현상에 영향을 미칠 수 있는 두 가지 조직적 행위자, 정부와 기업에

대해 서술하겠다.

제 2 절 정신질환 연구의 생의료화 (Biomedicalization) - 정부와 기업을 중심으로

정신질환의 유전화 현상은 유전적 위험인자에 근거하여 잠재적인 환자를 구성하고, 개인의 유전정보에 근거한 개별화된 치료 방식을 통해 정신질환을 관리해야 한다는 새로운 의학적 시선을 만들고 있다. 이는 정신질환을 진단하고 치료하는 의료 행동에 필요한 개념들이 새로운 방식으로 구성되고 있다는 것을 의미한다. 따라서, 정신질환의 유전화 현상에 대해 이해하기 위해서는 지식의 구성주의적 관점에 대해 논의할 필요가 있다. 예컨대, 쿤(1986)은 "패러다임"(paradigm)이라는 새로운 개념을 통해 과학의 발전 경로를 설명한다. 과학자들은 그 시대에 보편화된 방법론적, 이론적, 실험적 규칙과 틀 안에서 과학적인 문제를 정의하고 연구를 수행한다. 그는 이러한 틀과 규칙이 학계라는 지식 공동체 안에서의 토론과 합의를 통해 유지되고, 재구성되고, 전복될 수 있다고 봤다. 그의 관점에 따르면, 개별적인 사실들이 "과학"이라고 인정될 수 있는 이유는 그것들의 당시의 지식 공동체 안에서 합의된 관점에 근거하고 있기 때문이다.

라투르(2016)는 과학 지식에 대한 관계적 시각을 확장하여 어떻게 지식을 생산하는 연구자가 지식 바깥의 요소들과 결합하여 과학적 진리에 대한 지식 권력을 만들어내는지 논한다. 라투르의 관점에서 과학적 사실이 그것이 거짓임을 밝힐 수 있는 논쟁의 여지를 봉쇄할 수 있는 이유는 공통된 이해관계를 기반으로 과학적 사실을

둘러싸고 다양한 인적, 물적, 제도적 요소들이 촘촘한 네트워크를 구성하고 있기 때문이다. 예를 들어, 약물의 효과와 안전성을 증명하기 위해서 약물에 대한 전문적인 지식을 갖고 있는 과학자들, 약물을 개발하는 과정에 필요한 막대한 자본, 약물의 효과를 수치적으로 나타낼 수 있는 다양한 실험 장비들이 동원된다. 해당 지식은 약물을 의료 시장에 공급하여 자본을 축적하고자 하는 제약회사의 이해관계에 기반하고 있기 때문에, 이러한 다양한 요소들의 결합은 안정적으로 과학적 진리를 생산할 수 있다.

생의료화(biomedicalization) 개념은 과학의 구성주의적 관점과 의료화 논의를 결합한다. 클라크(2010)는 생의료화를 "복합적이고, 다양한 장소에서, 다양한 방향성을 가지고(complex, multisited, multidirectional) 진행되는 의료화 과정"이라고 정의한다. 다시 말해, 특정 문제가 의료적인 문제로 전환되는 의료화 과정이 사회적, 정치적, 과학기술적(technoscientific) 요소들이 결합되어 더욱 복잡한 방식으로 진행된다는 것이다 (Clarke et al., 2010; Conrad, 2005).

그는 생의료화를 촉진하는 다섯 가지 요소를 다음과 같이 열거한다. 첫째, 인간의 신체 및 건강을 상업화하는 의료 산업 복합체(medical industrial complex)는 생의료화로의 변화를 견인하는 대표적인 주체이다. 둘째, 개인의 건강을 위협하는 다양한 요소들이 계산되고 통제될 수 있기 됨에 따라서, 이러한 위험 요소들을 관리하는 의학적 관행이 널리 퍼졌다. 셋째, 기술 및 과학적 혁신은 의학적 질병의 범위를 확장하고 있다. 넷째, 많은 표본을 모집하여 특정 약물이나 치료법의 효과를 측정하는 무작위통제실험(randomized controlled trials)이 과학적인 지식을 생산하는 방식으로 보편화되었다. 다섯째, 환자들의 다양한 수요를 충족할 수 있는 맞춤형(customized) 치료법과 약물이 발전하고 있다.

전술한 다섯 가지 요소를 고려했을 때, 정신질환의 유전화는 생의료화의 대표적인 사례로 볼 수 있다. 유전학 연구들은 유전적 변이와 정신질환의 유의성을 밝힐 수 있는 많은 수의 표본 집단을 모아

대규모의 실험을 함으로써, 분자 단위의 신체적 기전에 개입하여 정신질환을 관리하고 통제할 수 있다는 의학적 근거들을 만들어내고 있다. 그리고, 의료 산업은 이러한 의학적 근거를 적극적으로 활용하여 개별적인 유전적 특성에 맞춘 치료법을 소비자들에게 제공하고 있다.

그렇다면, 유전자, 실험실, 기술적 요소들, 과학자 등의 물적, 인적 자원을 확보하고 이를 적재적소에 배치하여 유전자를 둘러싼 과학적 지식을 생산하는 주체는 누구일까? 연구자 개인의 학문적 호기심만으로 다양한 물적, 인적 자원이 동원되기는 쉽지 않다. 정신질환의 유전화 현상을 마음이라는 추상적인 신체의 기전조차 신체의 물리적, 화학적 단위를 통해 결정된다는 생물학적 정신의학의 주장이 보편화된 결과라고 보기도 어렵다. 결국, 현대 사회에서 다양한 물적, 인적 자원을 확보하고 이를 적재적소에 배치하고 관리할 수 있는 주체는 조직적 기관일 수밖에 없다. 정부, 기업, 학교 등의 다양한 기관들은 특정한 집단적인 이해관계를 위해 다양한 물적, 인적 요소를 동원하여 과학적 지식을 생산할 수 있다.

본 연구는 정신질환의 지식 구성에 영향을 미칠 수 있는 조직적 행위자로서 정부와 기업에 주목한다. 미국의 경우, 1970년대까지 주로 주정부가 운영하던 정신병원 시스템 (mental hospital system)은 정신질환을 치료하고 환자들을 관리하는 국가 정책의 근간이었다. 룰(2018)은 다음의 세 가지 원인에 의해 정신병원 시스템이 붕괴되고 연방정부(federal government)와 의료 산업(medical industry) 중심으로 정신질환을 둘러싼 의료 체계가 전환되었다고 설명한다. 첫째, 1973년 오일쇼크 이후의 10년 간의 저성장을 계기로 정신병원을 운영하는 재정적 비용에 대한 비판이 제기되었다. 둘째, 정신질환 환자가 격리용 시설에서 벗어나 시설 밖의 공동체에 통합될 수 있어야 한다는 사회적 담론이 발전하였다. 셋째, 환자가 자신이 받을 치료를 자율적으로 선택할 수 있어야 한다는 소비주의(consumerism)가 확산되었다. 연방정부는 병원 치료에 대한 재정적 지원을 줄이고 의료 산업의 발전을 유도함으로써 정신질환에 대한 재정적 부담을 줄이고자 했다. 이는 질병

관리를 둘러싸고, 정부의 재정적 효율성에 대한 이해관계와 사적 시장을 통해 이윤을 창출하려는 기업의 이해관계가 맞아떨어졌음을 의미한다.

그렇다면, 정부와 기업의 조직적 이해관계는 정신질환의 지식에 어떻게 반영될 수 있을까? 미국 정부 조직 내에서 정신질환 연구의 경향에 영향을 미칠 수 있는 대표적인 기관은 국립정신건강연구소(NIMH, National Institute of Mental Health)이다. 국립정신건강연구소는 정신질환의 연구에 있어 "생애주기에서 정신질환이 발병하는 경로를 분석"하고,"정신질환의 예방과 치료를 위해 노력하는 것" (NIMH, 2021, Strategic Plan Goals)을 전략적 목표로 삼고 있다. 즉, 정신질환의 위험 요소를 최대한 일찍 발견하는 것이 예방과 치료에 있어 최선의 결과를 낳을 수 있다는 것이다. 이러한 정부의 목표는 질환의 원인을 최대한 일찍 발견하여 병세가 악화되기 전에 질병을 최대한 효율적으로 관리하겠다는 정부의 경제적 동기를 보여준다. 정부의 목표는 연구자가 연구 주제를 설정하는 것에 있어 강한 영향을 미친다. 왜냐하면, 미국의 국립 보건원은 연구자들의 연구계획서를 받아 펀드를 제공하는데, 정부의 문제의식을 수용하지 않는다면 연구 기금을 받을 때 어려움이 있을 수 있기 때문이다 (조원광, 2017).

다른 한편, 정신질환에 대한 지식의 발전에 막대한 영향력을 행사한 사적 기관은 제약 및 바이오 기술 기업(pharmaceutical and biotechnology companies)이다. 실제로 정신질환이 완치될 수 있다고 믿게 된 것은 다양한 약물이 제약회사의 실험실에서 합성되기 시작한 이후였다 (Shorter, 2020). 정신약물학 연구는 화학 성분의 조합에 주목하는 제약회사와 생물의 세포 기전을 연구하는 바이오 기술 산업에 의해 주도되었다. 약물정신학의 발전은 의사들에게 약물을 공급하고 소비자들에게 약물의 효과를 설득하여 많은 이윤을 창출하고자 하는 기업의 이해관계가 있었기에 가능했다. 초기의 약물 개발은 주로 '뜻밖의 발견'에 의존했다. 정신질환에 효과를 보이는 발견이 이뤄지면 그것을 대규모의 실험을 통해서 통계적으로 그 약물의 효과를 입증하는

식이었다. 하지만, 분자생물학이 발전한 이후 질병과 연관된 생화학적 경로나 해부학적 구조에 대한 지식이 축적되었고, 이를 근거로 약물 개발에 필요한 물질을 전략적으로 선택하고 그 효과를 측정하는 것이 가능하게 되었다 (Busfield, 2006).

본 연구는 정부와 기업이 정신질환 연구의 유전화를 촉진하는 조직적 행위자가 될 수 있다고 본다. 정부는 아직 외현적인 증상이 보이지 않더라도 유전적 위험 인자에 근거하여 잠재적인 환자를 구성함으로써 정신질환을 조기에 효과적으로 관리할 수 있다. 기업은 개인의 유전적 특성에 근거하여 부작용을 줄인 맞춤형 약물을 개발하여 의료 시장을 확대할 수 있다. 그렇다면, 정부와 기업 중 어떤 조직적 주체가 정신질환 유전화를 주도하였을까? 나아가, 시간에 따라 각 주체의 영향력은 어떻게 변화하였을까? 본 연구는 구조적 토픽모델링 분석을 통해 해당 질문에 대한 답을 제시하고 그 함의를 서술하고자 한다.

본 연구는 조원광(2017)의 논문을 참조하여 전체적인 문제의식과 분석 방법론을 구성하였다. 그는 유방암을 다루는 전체 논문을 수집하여 정부의 펀드가 의료 지식에 미치는 영향을 측정하기 위해 구조적 토픽모델링 분석을 수행하였다. 그는 전체 논문을 정부의 펀드를 받은 논문과 그렇지 않은 논문으로 나누고, 정부의 펀드를 받은 논문의 경우 유방암 세포나 유전자 등의 분자적 대상에 주목하는 경향이 크다는 사실을 밝혔다. 본 연구는 정부의 펀드가 의료 지식에 미치는 영향력을 주목함과 동시에, 제약회사 및 바이오기술 회사와 연계된 논문을 따로 구분하여 기업의 영향력을 독립적으로 측정하고자 하였다. 그 이유는 두 가지로 정리할 수 있다. 첫째, 정신질환의 역사를 고려했을 때, 정부와 기업의 역할이 명확히 구분되기 때문이다. 정부의 목표는 질환을 효율적으로 관리하는 것이고, 기업은 질환을 약물로 치료하고자 한다. 둘째, 정신질환의 약물화 과정을 고려했을 때, 다양한 기업 중에서도 제약 및 바이오 기술 기업이 정신질환 연구의 유전화를 촉진할 수 있는 조직적 행위자이기 때문이다. 따라서, 본 연구는 두

주체의 영향력을 독립적으로 측정함으로써, 정신질환의 의료 지식을 구성함에 있어 두 주체의 역할이 어떻게 구분되고, 이러한 상이한 조직적 이해관계가 유전학이라는 특정 분야에서 어떻게 결합되는지 분석 결과 파트에서 서술하도록 하겠다.

제 3 장 자료 및 연구 방법

제 1 절 자료와 변수

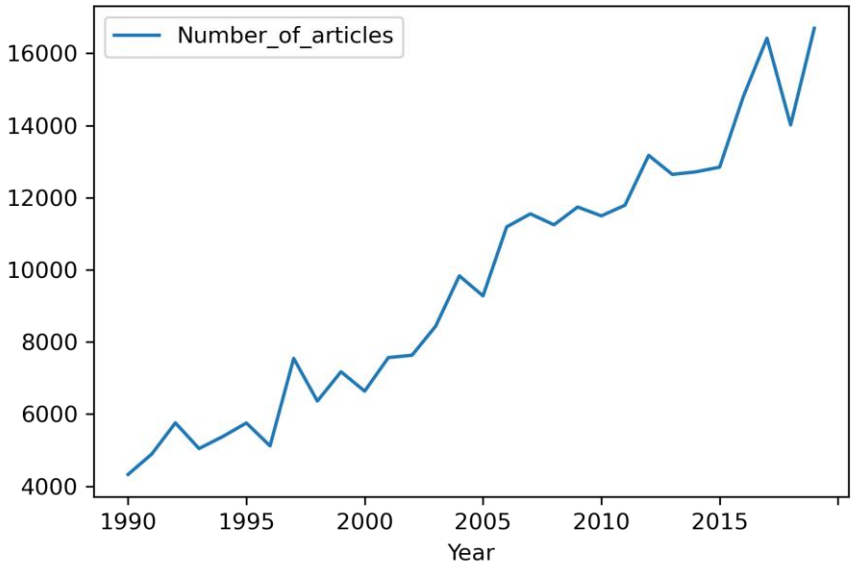
본 연구는 펍메드(PubMed) 데이터베이스에서 정신질환을 다룬 논문 정보를 수집하여 분석 데이터로 활용했다. 펍메드는 각종 생의학 및 생명과학 저널들과 온라인 서적에 있는 3천 4백만개 정도의 논문을 수록하고 있는 온라인 데이터베이스이다(PubMed, n.d.). 펍메드를 선택한 첫 번째 이유는 생의학(Biomedicine)[①] 분야에서 대표성 있는 데이터베이스라고 말할 수 있기 때문이다. 펍메드와 비견될 수 있는 생의학 데이터베이스로는 메드라인(MEDLINE)이 있다. 메드라인은 펍메드와 마찬가지로 미국 국립 보건원(National Institutes of Health, NIH) 소속의 의학 국립도서관(National Library of Medicine, NLM)에서 운영하는 데이터베이스로서, 생의학을 비롯한 생명과학 분야의 1천 9백만개 논문의 서지 정보를 수록하고 있다. 이 데이터베이스는 펍메드

① 다양한 방식으로 정의할 수 있지만, 생의학은 "신체의 구조와 기능, 이들의 정상과 비정상을 물리학, 화학, 생물학적 이해"(Lee, 1993)에 바탕을 둔 근대 의학을 말한다.

데이터베이스의 주요 요소이며, 많은 수의 MEDLINE 서지 정보는 PubMed에서도 열람할 수 있도록 링크가 기재되어 있다(Masic & Milinovic, 2012). 무엇보다, 펍메드는 일반 대중들도 전체 텍스트를 포함한 해당 논문의 모든 서지 정보를 무료로 활용할 수 있도록 허용했기 때문에, 좋은 접근성을 가지고 있다.

구체적인 논문 수집 기준을 열거하면 다음과 같다. 첫째, 펍메드는 의학 분야를 분류하기 위해 MeSH (Medical Subject Headings) 라는 인덱스를 제공하는데, 거기에서 Mental Disorders에 해당하는 논문을 추출하였다. 둘째, 인간을 대상하고 한 논문을 추출하였다. 셋째, 영어로 씌어진 논문만을 추출하였다. 영어는 학술 언어로서 보편화되어 있을 뿐만 아니라 텍스트 분석에 있어서 언어를 통일해야 하기 때문이다. 넷째, 초록이 있는 논문만을 추출하였다. 초록은 논문의 전체 내용을 압축적으로 제시하고 있기 때문에, 분석 대상을 초록에 있는 텍스트로 한정했을 때 보다 효율적인 통계적 분석을 수행할 수 있다. 다섯째, 1990년부터 2019년까지의 논문을 추출하였다. 그 이유는 1980년대까지는 논문의 메타데이터가 잘 구축되어 있지 않은 경우가 많기 때문이다. 여섯째, 미국에서 발행된 논문만을 추출했다. 미국은 정신질환의 의료화를 주도한 대표적인 국가이고, 전체 논문 중에서 미국에서 발행된 논문의 비중이 크다. 일곱째, 기업에 연계된 논문과 그렇지 않은 논문들을 구분하기 위해 저자들의 소속이 기재된 논문만을 추출했다. <그림 1>은 이와 같은 조건을 만족하는 전체 논문의 증가 추이를 보여준다.

그림 1 전체 논문의 수 추이



펍메드는 논문의 텍스트뿐만 아니라, 다양한 메타데이터를 제공하고 있다. 저자는 정부와 기업이 정신질환 연구의 담론에 끼치는 영향을 측정하기 위해 이러한 메타데이터들을 활용하여 두 가지 더미 변수를 구성하였다. 첫째, 정부의 후원을 받는 논문과 그렇지 않은 논문을 구분하였다. 이를 위해, Grant Number 메타데이터에서 NIH 소속의 건강 기관들(health agencies)의 이름이 포함된 논문들은 연방 정부의 후원을 받는 논문으로 분류하였다.

둘째, 저자들의 소속이 기재된 Affiliation(AD) 메타데이터를 활용하여 기업과 연계된 논문과 그렇지 않은 논문을 구분하였다. 에반스(2010)는 논문의 저자 목록과 사사문구(acknowledgement) 메타데이터를 기반으로 특정 논문이 기업과의 협업 아래 씌어진 논문과 그렇지 않은 논문을 구분한다. 구체적으로, 그는 기업과 학계가 공동으로 연구를 수행하는 협력적 관계, 그리고 연구 결과를 활용할 권리를 얻는 대신 연구할 자금을 제공하는 계약적 관계, 두 가지로 산학 협력 관계를 구분하여 정의한다. 에반스의 방식을 따라, 본 연구에서는

우선 기업 소속의 과학자가 저자 목록에 포함되어 있다면, 해당 논문이 기업의 이해관계를 반영할 가능성이 높다고 간주하고 기업과 연계된 논문으로 분류하였다. Affiliation 메타데이터는 저자들이 소속된 회사(혹은 학교), 부서, 위치 등을 기재하고 있다. Statista(2019)에서 약 판매량, R&D 투자액, 시장 가치를 기준으로 상위 40개 제약회사(pharmaceutical company) 및 바이오 기술 회사(biotechnology company)를 선정하고, 이들 이름이 Affiliation 메타데이터에 포함되었다면 기업과 연계된 논문으로 구분하였다. <부록 1>은 해당 변수를 구성할 때 활용한 제약 회사 및 바이오 기술 회사 목록 보여준다. 전술한 데이터 처리 과정을 통해, 본 연구는 정부 펀드를 받은 논문 81,374 개, 기업과 연계된 논문 10,425 개를 추출하였다. 두 변수에 따라 데이터를 분류했을 때 나오는 행렬을 <표 1>에 정리하였다.

표 1 변수에 따른 데이터 분류 행렬

	기업 소속 논문	그 외	합계
정부의 후원을 받는 논문	3,416	77,958	81,374
그 외	7,009	200,660	207,669
합계	10,425	278,618	289,043

더불어, 전체 논문 중 Conflict of Interest Statement(COIS) 메타데이터를 참고하였다. 이 메타데이터는 저자가 특정한 이해관계를 반영하고 있을 경우 이를 명시하도록 한다. 본 연구는 이 메타데이터가 표기되어있는 2,608개의 논문을 하나씩 검토하여 컨설팅, 데이터 공유, 연구비 지원 등의 형태로 제약 및 바이오기술 회사와의 협력 및 계약이

명시되어있는 경우 따로 구분하였다. 해당 조건을 만족하는 논문의 개수는 201개였다. 이들 논문과 앞서 구분한 10,425개의 논문과 이들 논문을 합치면 기업과 연계된 논문은 10,598개로 증가한다. 이러한 구분법으로 돌린 모델은 <부록 2>에 간단히 제시하였다.

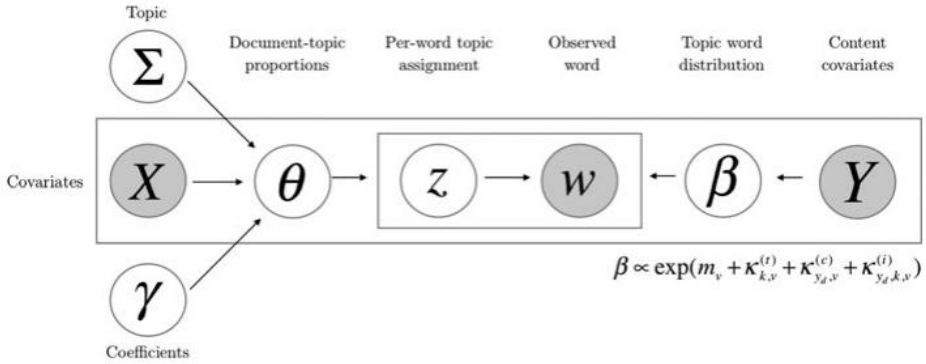
제 2 절 구조적 토픽모델링

본 연구는 정신 질환 분야 내에서 유전학 연구들을 추출하고, 그것이 정부 및 기업에 의해 어떻게 구성되는지 파악하기 위해 구조적 토픽모델링 (Structural Topic Modeling, STM)을 활용한다. 토픽모델링은 대량의 언어로 구성된 문서들 안에서 담론적 구조를 추출하기 위해 광범위하게 사용되고 있는 기계학습 방법이다. 토픽모델링은 사전에 지정된 토픽의 개수에 따라 토픽을 추출하는데, 이때 토픽은 각기 다른 확률분포를 보이는 단어들의 집합이 된다. 예를 들어, 한 토픽에서 중요 단어로 추출된 단어들은 임의의 단어 분포에 비해 문서들 안에서 동시에 출현할 확률이 높은 단어들이다. 따라서, 토픽 모델링은 단어들 간의 관계에서 생성되는 의미를 파악할 수 있다는 점에서 장점을 갖고 있다. 예를 들어, “drug”라는 단어는 “abuse”, “prisoner”, “race”와 같은 단어들과 묶일 때와 “patient”, “safety”, “effectiveness”와 같은 단어들과 묶을 때 완전히 다른 의미를 갖게 될 것이다.

구조적 토픽모델링은 “문서의 메타데이터와 상관관계를 갖는 문서-토픽 및 토픽-단어 분포를 통해 문서를 생성하는”(Roberts, Stewart, & Tingley, 2019) 모델로 정의할 수 있다. 여기서는 모델을 구성하는 수식을 자세히 다루기보다 로버트(2016)의 논의를 참고하여 구조적 토픽모델링의 생성 과정(generative process)을 간단히 요약하는

방식으로 서술하겠다. 첫째, 토픽 발현 모델(topic prevalence model)은 한 문서에 있는 단어들이 어떻게 한 토픽에 할당되는지를 통제한다. 이 때, 문서 내에서 토픽의 비율(document-topic proportions, Θ)은 문서 수준의 공변량(X), 즉 문서의 메타데이터에 따라 변화한다. 둘째, 토픽 내용 모델(topical content model)은 한 토픽 안의 단어들의 분포를 통제한다. 토픽($\kappa^{(t)}_{k,v}$), 공변량($\kappa^{(c)}_{yd\ ,v}$), 그리고 토픽-공변량 상호작용항($\kappa^{(i)}_{yd,k,v}$)이 토픽 내에서의 단어 분포(topic word distribution, β)를 결정한다. 셋째, 핵심 언어 모델(core language model)은 이 두 가지 모델을 결합하여 각각의 문서에서의 실제 단어들을 생성한다. 문서-토픽 비율 벡터가 주어지면, 다항분포 Z(per-word topic assignment)으로부터 문서 안에 있는 각 단어 당 토픽이 추출된다. 단어 당 토픽이 할당되면, 각 토픽을 구성하는 단어들의 분포로부터 실제 단어(w, observed word)가 선택된다. 여기서 중요한 것은, 문서-토픽 비율(Θ)과 토픽의 단어 분포(β)가 모든 문서에 공유되는 사전 파라미터에 의해 생성되는 것이 아니라, 문서 수준의 공변량과 상관관계를 갖는다는 것이다. <그림 2>는 구조적 토픽모델의 생성과정을 간단한 그래픽으로 보여준다.

그림 2 구조적 토픽모델의 생성과정 (Roberts et al., 2016)



따라서, 구조적 토픽모델링은 문서의 메타데이터 정보와 토픽 발현 및 토픽 내용과의 관계성을 분석할 수 있다는 점에서 이점이 있다 (Roberts et al., 2019; 김란우 & 송수연, 2020). 구조적 토픽모델링은 시간과 같은 연속적인 공변량(covariates)을 모델에 포함시킬 수도 있다. 이때, 토픽의 발현 확률과 시간은 서로 비선형 관계를 갖기 때문에, 일반 선형 회귀모형의 계수를 비선형 함수로 대체하여 표현하는 일반화가법모형 (Generalized Additive Models, GAM)을 활용하여 평탄화(smoothing) 과정을 수행한다 (James, Witten, Hastie, & Tibshirani, 2013; 송준모 & 강정한, 2018). 시간 변수는 토픽 모형의 결과가 연도에 따라 편향되지 않도록 하는 효과도 있다 (김란우 & 송수연, 2020). <그림 1>에서 본 것처럼 정신질환 연구는 최근 급성장하였기 때문에, 연도를 통제하지 않고 모형을 적용할 경우 최근의 담론을 이전의 담론에 비해 더 반영할 가능성이 있다. 하지만, 구조적 토픽모델링은 연도별 샘플 수에 따라 논문에 가중치를 다르게 부여함으로서 모든 연도의 토픽이 고르게 결과에 반영되도록 도와준다 (김란우 & 송수연, 2020). 정리하면, 본 연구는 앞서 언급한 기업/비기업, 정부/비정부 변수와 함께 연도 변수를 공변량에 포함하고, 이들이 토픽 발현에 어떻게 영향을 미치는지 규명하는 것에 집중한다.

본 연구에서는 구조적 토픽 모델링을 통해 추출된 토픽들을 학계의 담론 혹은 정신질환 연구를 구성하는 세부 연구 분야로 상정하였다. 즉, 본 연구에서 연구 대상으로 삼은 정신질환 연구는 구조적 토픽 모델링에 의해 추출된 세부 연구 분야들에 의해 구성되어 있다고 볼 수 있다. 토픽모델링은 각 문서를 한 토픽에 배타적으로 분류하는 것이 아니라 각 문서를 여러 토픽의 합으로 간주한다. 따라서, 한 편의 논문이 가지는 토픽 비율의 합은 언제나 1이 된다 (Roberts et al., 2019). 대부분의 학술적인 논문은 논리적으로 일관적이면서 해당 논문에만 고유한 개념들로 구성되는 것이 아니라 다른 논문에서도 사용된 개념들이 새로운 문제의식 아래 논리적으로 결합하면서 만들어진다. 특히, 유전학, 약물학, 심리학 등 다양한 학문적 분야가

융합하고 있는 정신질환 연구의 특성을 고려하였을 때, 이러한 토픽모델링의 특성은 정신질환 논문의 언어 구조를 분석할 때 큰 장점이 될 수 있다.

제 3 절 전처리 및 통계 도구

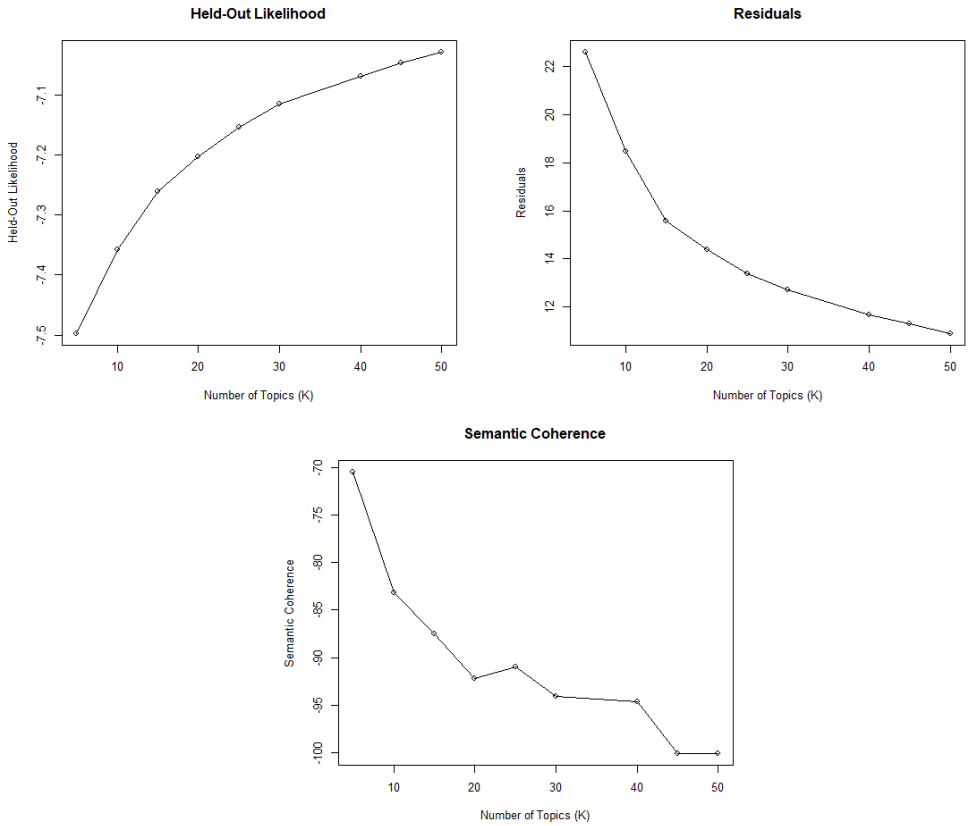
본 연구는 데이터 다운로드와 전처리는 파이썬 (Python)을 활용하였고, 통계분석은 알 (R) 프로그램을 활용하였다. 우선, BioPython 패키지에 있는 Entrez 모듈을 통해, 총 4개의 조건 – “Mental Disorders 논문”, “인간을 대상으로 한 논문”, “영어로 쓰여진 논문”, “초록이 존재하는 논문” – 을 만족하는 논문 총 838,253 개의 데이터를 다운로드 받았다. 그 중에서, 나머지 3개의 조건 – “1990년부터 2019년까지의 논문”, “미국에서 발행된 논문”, “저자의 소속이 기재된 논문” – 을 만족하는 논문 289,043개를 추출하였다. 보다 효율적인 통계분석을 위해, 전체 문서를 소문자로 전환하고, 구두점, 숫자, 중복 여백, 불용어들을 제거했다. 표제어 추출(lemmatization) 과정을 거쳐, 서로 다른 형태의 단어들을 어근 단어로 통일하였다. 마지막으로, 10개 이하의 문서에서 발견되는 단어들을 제거하였다. 그 결과, 총 126,870 개의 단어 중에 99253개를 제거하여, 27,617개의 단어를 분석 대상을 활용하였다. 구조적 토픽모델링 알고리즘은 로버츠, 스튜어츠, 팅리(2019)가 제작한 구조적 토픽모델링 알 패키지(R package)의 1.3.5 버전을 활용하였다.

제 4 장 분석 결과

본 연구는 전체 문서에서 총 20개의 토픽을 추출하였다. 구조적 토픽모델링은 모형의 성능을 평가해줄 비교대상인 훈련 집단(training set)과 시험 집단(test set)이 존재하지 않는 비지도 기계학습(unsupervised machine learning)이기 때문에, 토픽의 수를 결정하는 것은 매우 어려운 일이다 (Grimmer & Stewart, 2013). 하지만, R의 stm 패키지에서 제공하는 진단통계치들을 참조해볼 수는 있다. <그림 2>를 보면, 토픽의 개수가 증가할수록 held-out 가능도(held-out likelihood)와 잔차(residual)는 개선되는 반면, 의미론적 일관성(semantic coherence)이 감소한다. [②] 즉,“토픽의 개수가 늘어날수록 토픽 사이의 경계가 뚜렷해지며 모형의 성능이 개선되지만, 의미가 과도하게 세분화되어 범주로서의 가치가 없어지는 문제” (송준모 & 강정한, 2018)가 발생한다. 따라서, 이상적인 토픽의 수를 결정하는 데에는 연구자의 선택이 필요하다. 본 연구는 여러 개의 토픽 수를 비교해봤을 때, 20개가 정신의학의 세부 분야를 드러낼 수 있는 이상적인 토픽 수라고 판단했다.

[②] 토픽모델링에서의 held-out 가능도는 특정 단어를 한 문서에서 제거하여 새로운 데이터셋을 구성하고, 모델을 새로운 데이터셋에 돌렸을 때 그 단어가 해당 문서에서 등장할 확률을 말한다. 특정 단어를 제거하였음에도 그 단어가 해당 문서에 등장한다는 것을 높은 확률로 추정할 수 있다면 전체 텍스트의 단어 분포를 잘 반영한 좋은 성능의 모델이라고 볼 수 있다. 이는 전체 데이터 중 일부분을 떼낸 데이터에 통계 모델을 적용하여 모델의 적합성을 판단하는 교차타당도(cross-validation) 개념과 유사하다. 잔차가 높다는 것은 잔차의 분산이 크다는 것을 의미하고, 토픽의 수를 늘려 모델의 설명력을 높인다면 잔차를 줄일 수 있다. 의미론적 일관성은 주요 단어들이 자주 함께 등장할 때 극대화된다(Roberts et al., 2019). 만약, 토픽의 수가 적다면 주요 단어들이 자주 함께 등장할 확률이 높아질 것이다. 하지만, 자주 함께 등장하는 단어들일지라도 서로 잘 공유되지 않는 다른 단어와 각각 연결되어 있다면 한 토픽 내에서도 의미론적 차이가 나타날 수 있다.

그림 3 진단통계치

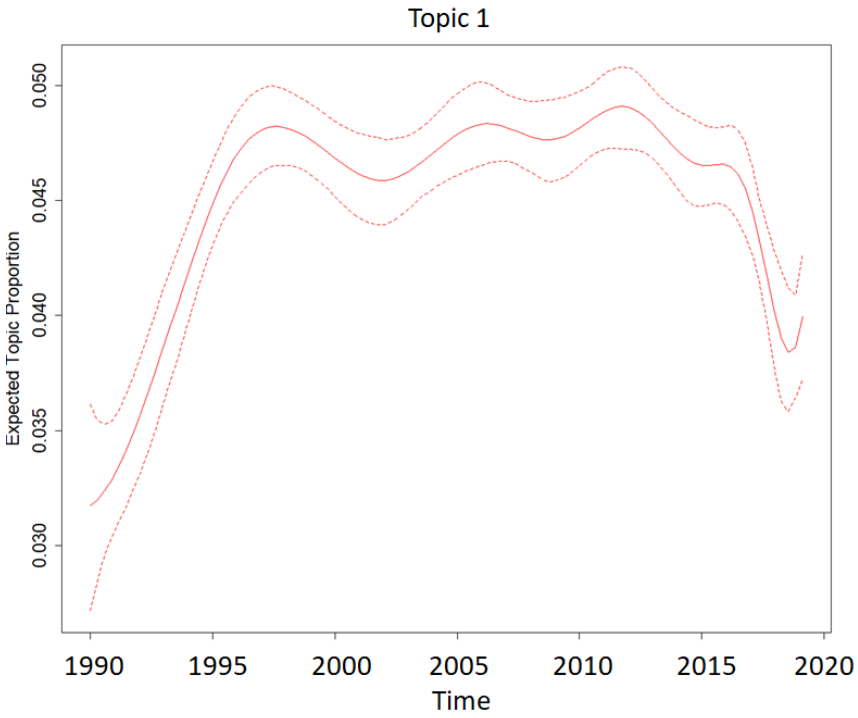


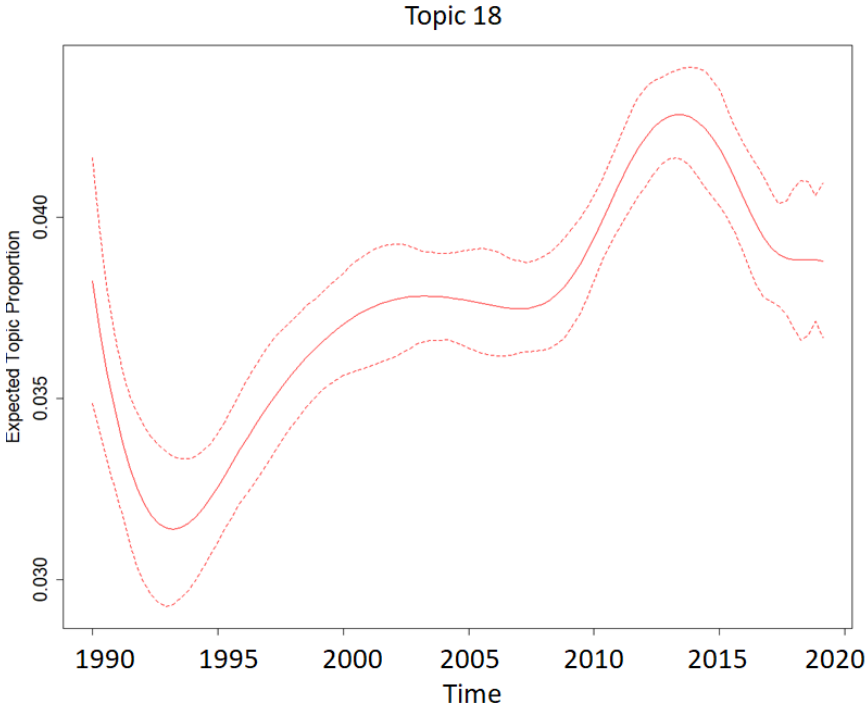
저자는 로버트(2019)가 제시한 두 접근법에 따라 토픽을 해석하였다. 첫째, 각 토픽에서 가장 높은 확률로 추출된 단어(highest probability words)와 FREX 기법에 의해 추출된 단어 각각 10개를 고려하였다. FREX 기법은 특정한 토픽에 배타적으로 분포되어 있는지(exclusivity), 그리고 얼마나 많이 분포되어 있는지(frequency)를 함께 고려하여 각 토픽을 대표하는 중요 단어를 추출한다 (Roberts et al., 2019). <부록 3>은 각 토픽마다 어떤 단어들이 추출되었는가 보여준다. 둘째, 각각의 토픽의 비중이 가장 높은 다섯 개의 문서들을 추출하여 검토하였다. 토픽모델링에서 각 텍스트는 모든 토픽의 합으로 계산된다. 따라서, 특정 토픽의 비중이 매우 높은 텍스트라면 해당 토픽을 대표하는 텍스트라고 봐도 무방할 것이다. 본 연구에서는 20개의

토픽을 뽑았으므로 총 100개의 논문의 초록을 검토하였고, 그 해석 결과를 바탕으로 각 토픽의 제목을 정하였다. <부록 4>는 각 토픽의 제목을 보여준다.

각 토픽의 제목과 주요 단어들을 살펴보면, 이들을 정신질환을 둘러싼 의료 행동의 종류에 따라 몇몇 카테고리로 나눠서 설명해볼 수 있다. 첫째, 정신질환의 물리적 병인으로서 유전자와 뇌에 대한 토픽이 존재한다. 1번 토픽은 정신질환과 유전자의 연관성을 다루고 있다. <그림 4>에서 위쪽 그림은 시간에 따라 1번 토픽의 발현 확률이 어떻게 변화하는지 보여준다. 이를 보면, 게놈 프로젝트가 활발히 진행되던 때인 1990년대 초반에 토픽의 발현 확률이 급격하게 증가한 후, 그 수치가 일정하게 계속 유지되고 있다는 것을 알 수 있다. 이는 정신의학과 유전학의 학제 간 연구가 단기적인 유행이 아니라 정신질환 연구 내에서 세부 분야로서 제도화되었음을 보여준다. 뇌를 다루는 대표적인 토픽은 8번, 15번, 18번 토픽으로 분산되어 있다. 8번과 15번 토픽은 치매 및 알츠하이머 질환에 주목하는 반면, 18번 토픽은 뇌 신경의 전달물질과 정신질환의 연관성에 주목한다. 따라서, 18번 토픽이 정신질환의 물리적 병인으로서 뇌를 보다 고유하게 다룬다고 볼 수 있다. <그림 4> 아래 그림을 보면, 18번 토픽 또한 시간에 따라 꾸준히 증가하고 있다. 이를 통해, 유전학(Genetics)과 신경학(Neurology)이 정신의학과 융합하면서, 정신질환의 생물학적 원인을 탐구하는 학문적 경향이 정신질환 연구 내에서 그 영향력을 확장하고 있다는 것을 알 수 있다.

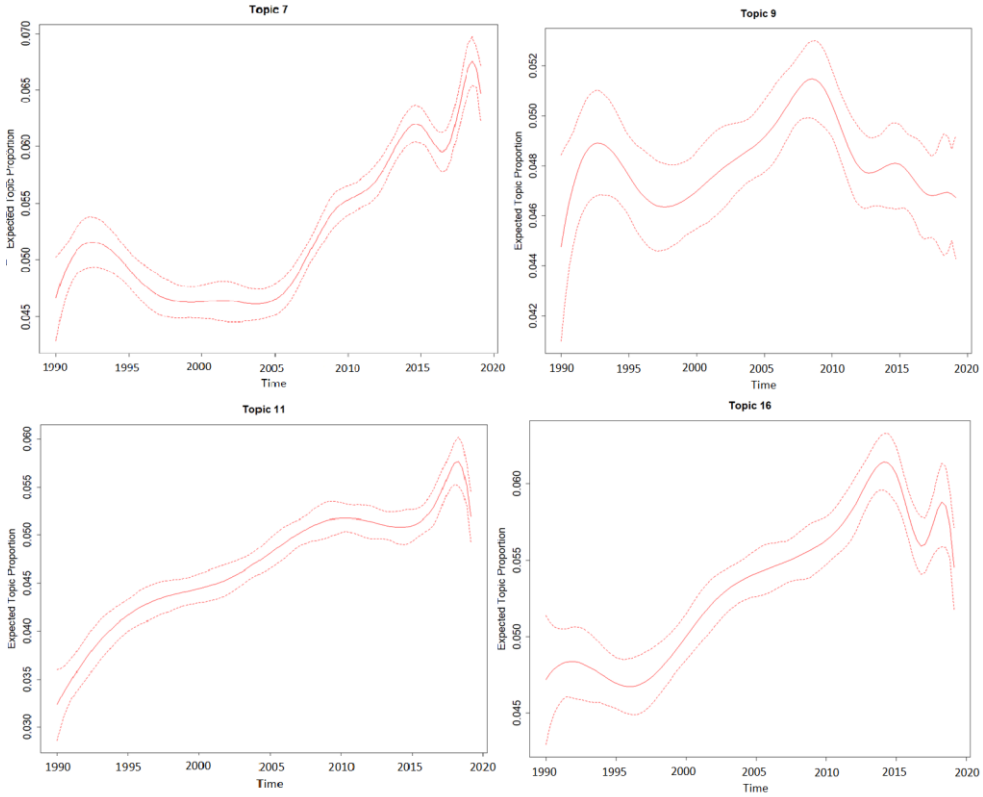
그림 4 1, 18번 토픽의 출간 시기에 따른 발현 확률의 변화





둘째, 정신질환을 그 신체적, 정신적 증상에 따라 분류하고 이를 구체적으로 기술하는 토픽이 검출되었다. 구체적으로, 불안 및 우울 (4번 토픽), 감정 및 행동 장애 (7번 토픽), ADHD (9번 토픽), 섭식 장애 (10번 토픽), PTSD (11번 토픽), 언어 장애 (16번 토픽), 수면 장애 (17번 토픽), 흡연, 음주, 약물 중독 (19번 토픽) 등 DSM에 있는 정신질환이 다뤄지고 있다. <그림 5>를 보면, 이들 토픽의 발현확률을 시간에 따라 관찰했을 때, 사회적 관계와 영향을 주고 받는 정신질환의 증상을 다루는 토픽들 (7, 9, 11, 16번 토픽)의 비율이 증가하고 있는 것을 확인할 수 있다. 감정, 행동, 언어 장애는 사회적 관계에서 비롯되기도 하며, 향후 사회적 관계를 형성할 때 많은 영향을 미치기도 한다. ADHD를 다루는 9번 토픽의 경우, 주로 ADHD가 아동의 친교 관계나 학업에 있어 어떤 영향을 미치는지 주목하였다. PTSD를 다루는 11번 토픽은 PTSD 환자들이 어떻게 일상 생활을 회복할 수 있는지를 주로 다루고 있다.

그림 5 7, 9, 11, 16번 토픽의 출간 시기에 따른 발현 확률의 변화



셋째, 정신질환 환자들을 위한 약물 치료의 효과 및 안전성에 주목하는 토픽들이 검출되었다. 3번 토픽은 무작위 통제실험(randomized controlled trial)을 통해 약물의 효과와 안전성에 대해 다루고 있다.

정부와 기업은 정신질환에 대한 지식을 어떻게 구성하고 있을까? 정부/비정부 및 기업/비기업 더미 변수와 토픽 발현 확률의 관계를 봤을 때, 다음과 같은 통계 결과를 확인할 수 있다. <그림 6>와 <그림 7>은 정부/비정부 변수에 따라 토픽의 발현 확률이 어떻게 분포하는지 보여준다. 특정 토픽이 오른쪽에 위치할 수록, 해당 토픽은 정부의 펀드를 받은 논문에 우세하게 나타난다는 것을 의미하고, 왼쪽에 위치할 수록 정부의 펀드를 받지 않은 논문에 우세하게 나타난다는 것을 의미한다. 기업/비기업 변수의 경우에도 마찬가지다. 특정 토픽이 오른

쪽에 위치할 수록 기업 소속의 저자가 있는 논문에 우세하게 나타난다는 것을 의미하고, 왼쪽에 위치할 수록 그렇지 않은 논문에 우세하게 나타난다는 것을 의미한다. 1, 2, 8, 9, 11, 15, 16, 18, 19, 20번 토픽은 정부의 펀드를 받은 논문에서 우세하게 나타나고, 3, 6, 8, 15, 18, 20번 토픽은 기업과 연계된 논문에서 우세하게 나타난다. 6, 7번 그림에서 토픽 별로 가장 높은 확률로 추출된 단어 세 개씩 표시하여 각 토픽의 내용을 쉽게 확인할 수 있도록 하였다.

그림 6 정부/비정부 변수와 토픽 발현 확률의 관계

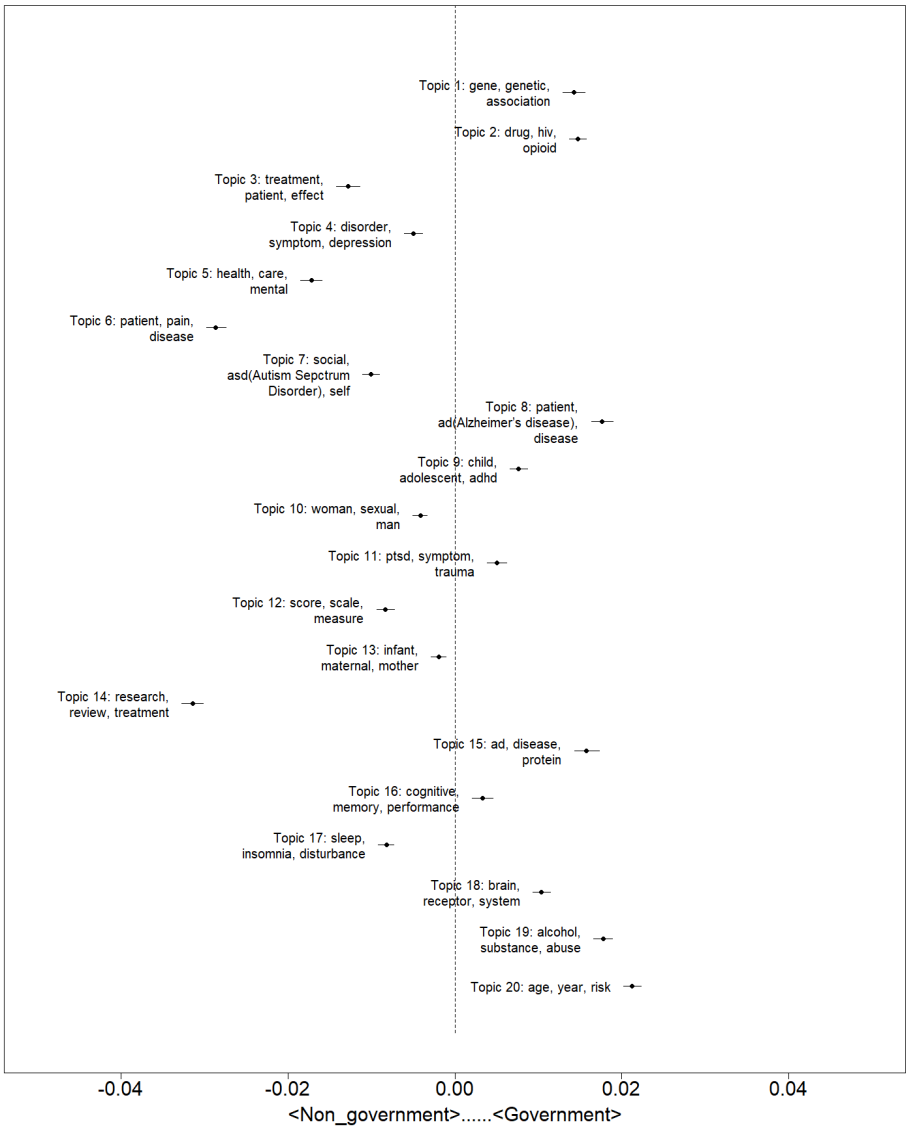
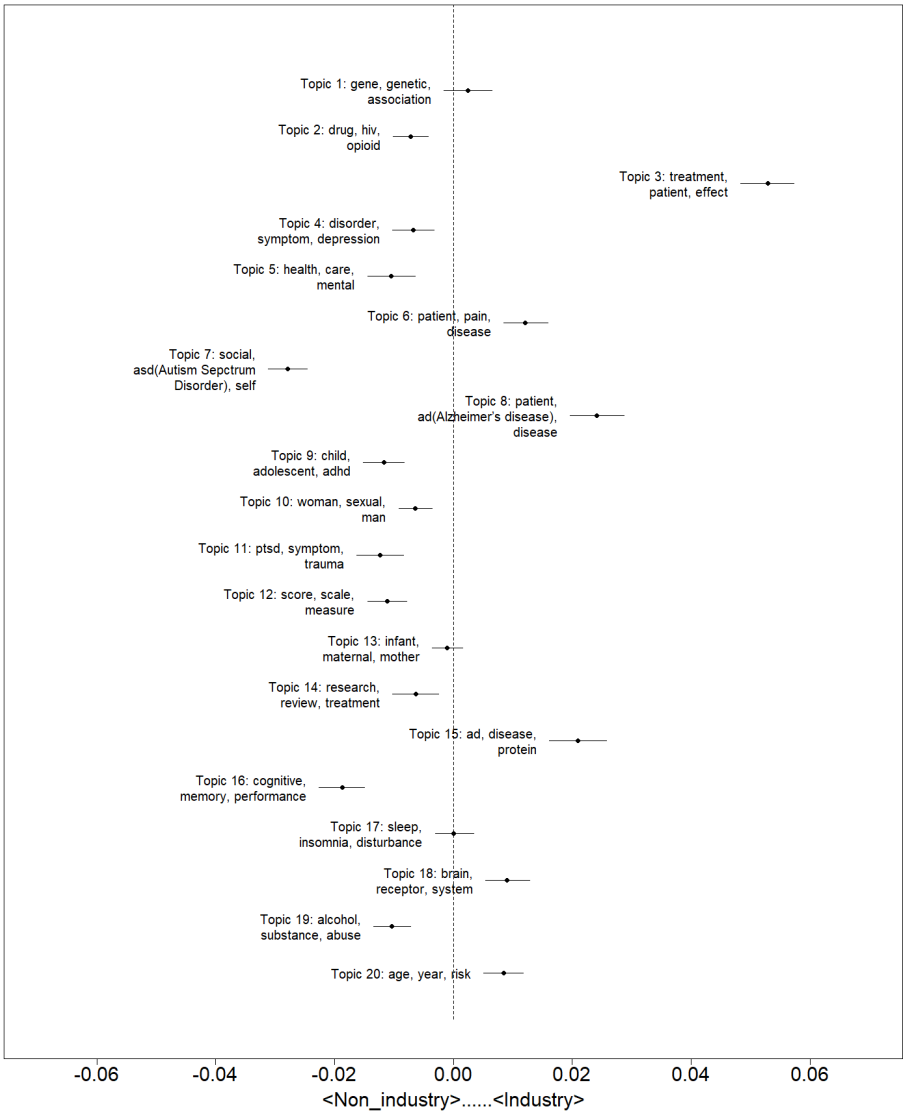


그림 7 기업/비기업 변수와 토픽 발현 확률의 관계



앞서 정리한 세 가지 카테고리와 연결시켜 봤을 때, 다음과 같은 결과를 이끌어 낼 수 있다. 첫째, 정부는 뇌와 유전자로 대표되는

정신질환의 생물학적 원인에 주목하는 경향이 있다. 구체적으로, 정부는 뇌와 유전자로 대표되는 정신질환의 생물학적 원인을 파악하는 것에 집중한다. 이는 정신질환의 위험 요소를 일찍 발견하여 질환을 최대한 효율적으로 관리하겠다는 정부의 이해관계를 반영한다. 특히, 1번, 18번 토픽의 발현 확률이 증가하고 있는 경향을 고려했을 때, 정부가 생물학적 정신의학의 발전을 추동하고 있다는 것을 알 수 있다. 예컨대, 유럽에서 시작되어 세계 각지로 전파된 바이오은행 프로젝트(biobank project)는 개인의 생물학적 정보를 수집하여 인구의 건강을 관리하려는 정부의 이해관계를 보여주는 대표적인 사례이다. 바이오은행 프로젝트는 혈액, 소변, 타액 샘플을 인구 코호트에 따라 대규모로 수집하여 유전 및 건강 정보를 축적하는 작업으로서, 질환의 생물학적 원인을 파악하는 작업에 중요한 자료를 제공하고 있다 (Claussnitzer et al., 2020). 다양한 기업들도 바이오은행의 주요 투자자로 참여하고 있기는 하지만, 바이오은행의 핵심적인 자금 제공자는 정부 기관 및 비영리 기관이다.

둘째, 정부는 사회적 관계와 연관된 정신질환의 증상에 주목하는 경향이 있다. 7, 9, 11, 16번 토픽 중에서 9, 11, 16번 토픽이 정부의 펀드를 받은 논문에서 우세하게 나타난다. 즉, 정부는 증상을 기준으로 환자 집단을 구성하고, 치료를 통해 그들이 사회적 관계를 회복하는 것에 관심을 둔다고 볼 수 있다. 그 이유는 무엇일까? 국가는 고도의 의료 기술을 통해 질병을 예방하고 치료하는 것을 넘어 주거, 영양, 위생, 지역사회 등 다양한 영역과 연관하여 건강 및 질병 문제에 대한 종합적인 해결을 도모한다 (조병희, 2006). 이것은 시장이 해결할 수 없는 영역이다. 다시 말해, 정부는 단순히 질병의 제거를 넘어 사람들이 높은 삶의 질을 통해 건강한 생활을 유지하는 것에 초점을 맞춘다. 예컨대, 행동 장애, 외상후 스트레스 장애, 언어 장애 등은 단순히 개인의 건강 차원의 문제가 아니라, 개인이 사회적 관계를 유지하는 것에 있어 필수적인 신체적, 정신적 능력에 부정적인 영향을 줄 수 있는 질환이다. 국가의 입장에서 이러한 질환의 증상과 환자의 고통을 파악하고 그들이 안정적으로 사회에 통합될 수 있도록 돕는 것은 중요한

목표가 될 수 있다.

셋째, 정부가 정신질환의 생물학적 원인과 신체적 증상에 관심을 갖는다면, 기업은 정신질환의 의료적 치료에 관심을 갖는다. 약물의 안전성과 효과를 다루는 3번 토픽은 기업과 연관된 논문에서 우세하게 나타난다. 약물 치료에 대한 과학적 지식을 생산하여 약물의 소비층을 늘리고자 하는 기업의 이해관계를 고려했을 때, 이는 합당한 결과라고 볼 수 있다.

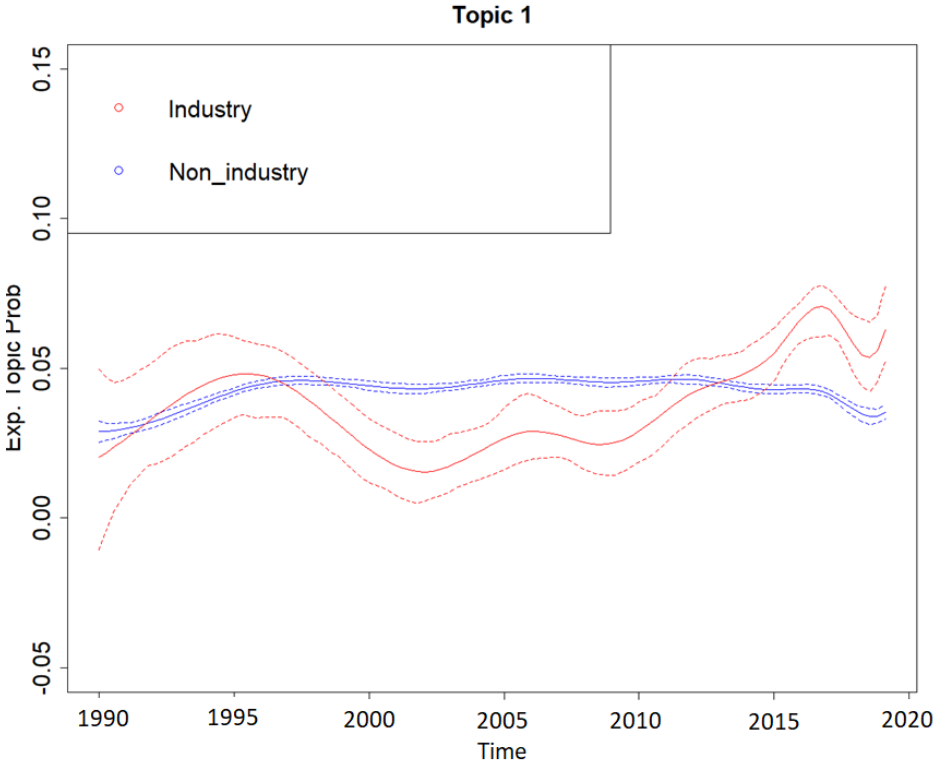
정리하면, 정부와 기업은 정신질환에 대한 과학적 지식을 생산하는 과정에서 일종의 역할 분담을 하고 있다. 정부가 원인과 증상에 집중한다면, 기업은 치료에 집중한다. 그렇다면, 정신질환의 진단, 예측, 치료에 대한 포괄적인 의학적 근거를 축적하고 있는 유전학 연구는 어떤 조직 행위자에 의해 주도될까? 정신질환과 유전자의 연관성을 다루는 1번 토픽은 정부의 펀드를 받은 논문에서 우세하게 나타난다. 하지만, <그림 7>을 보면, 기업/비기업 변수와 출간 시기 변수의 상호작용을 고려했을 때, 기업 소속의 논문에서의 토픽 발현 확률은 2010년대 들어 꾸준히 증가하는 반면, 그렇지 않은 논문에서의 토픽 발현 확률은 거의 같은 수준에서 유지되고 있는 것을 볼 수 있다.

그렇다면, 왜 2010년대를 기점으로 기업 소속 논문의 토픽 발현 확률이 증가했을까? 1990년대와 2000년대에 걸쳐 유전자 염기서열(genetic sequences)에 대한 지식을 기반으로 특정 단백질을 겨냥하여 약물적 효과를 내는 미세-분자 약물(small molecule drugs)에 대한 연구가 활발하게 진행되었다. 이러한 약물유전학의 발전은 인간 유전자에 대한 광범위한 지식이 축적되었기에 가능했다 (Nunnally, Webster, Brown, & Cohen, 2005). 이는 유전학이 정부와 기업의 이해관계가 접목될 수 있는 분야라는 것은 보여준다. 정부는 유전학 연구를 통해 정신질환의 생물학적 원인을 파악할 수 있고, 기업은 환자의 유전적 특성을 고려한 약물을 통해 부작용을 최소화하고 정신질환에 대한 유전자의 신체적 작용을 억제할 수 있다. 본 통계 결과는 정부의 펀드를 통해 발전한 유전학 연구의 성과가 점차 기업이

주도하는 실질적인 치료의 영역으로 확장되고 있다는 사실을 보여준다.[③]

[③] 여기에 한 가지 이유를 더 추측해볼 수 있다. 2013년 미국 대법원은 인간의 유전자는 특허(patent)의 대상이 될 수 없다는 결정을 내렸다. 대법원은 이와 같은 결정의 근거로 인간의 유전자는 자연의 산물(product of nature)이라는 주장을 내세웠다. 인간 유전자를 발견한다고 새로운 무언가가 창조될 수 있는 것이 아니기 때문에 보호될 수 있는 지적 재산으로 인정할 수 없다는 것이다. 대법원의 결정이 내려지기 전까지 4,300개 정도의 특허가 인간 유전자에 승인되었다 (NIH, 2021). 2009년을 기준으로, 유전자에 대해 특허를 받은 상위 30개 기관 중 20개가 제약 및 바이오기술 기업이었다 (Cook-Deegan & Heaney, 2010). 특허가 승인되면 그 보유자는 유전자 실험, 연구 등의 상업적, 학문적 활동에 대해 해당 유전자가 어떻게 활용될 수 있는지 20년간 결정할 수 있는 권리가 주어졌다. 만약 특허가 승인된 유전자의 다수가 상업적 가치가 큰 정보를 담고 있을 가능성이 크다고 가정한다면, 2013년 이후 인간의 유전자를 상업적으로 활용할 수 있는 문턱이 보다 낮아졌다고 말할 수 있다. 하지만, 실제로 특허가 승인되었던 유전자에 대한 접근성이 높아졌는지 말하기 위해서는 연구 논문에서 활용된 유전자 종류를 추출해보는 작업이 필요할 것이다. 유전자에 대한 지적 재산권의 존폐가 연구 및 개발의 혁신과 양적 성장에 어떤 영향을 미칠 것인가에 대해 많은 논쟁이 제기되었기 때문에 (Cook-Deegan & Heaney, 2010; Heller & Eisenberg, 1998; Nunnally et al., 2005), 이는 향후 연구를 통해 규명될 필요가 있다.

그림 8 1번 토픽에서 기업/비기업 변수와 출간 시기의 상호작용 효과



제 5 장 마치며

정신질환의 유전화 현상은 유전적 위험인자에 근거하여 잠재적인 환자를 구성하고, 개인의 유전정보에 근거한 치료 방식을 통해 정신질환을 관리해야 한다는 새로운 의학적 시선을 만들고 있다. 어떻게 유전학 연구는 정신질환의 진단, 예측, 치료에 대한 포괄적인 의학적

근거를 축적할 수 있었을까? 유전학 연구에 필요한 다양한 기술적, 인적, 물적 자원을 확보하고 이를 적재적소에 배치하여 유전자를 둘러싼 과학적 지식을 생산하는 주체는 누구일까? 본 연구는 해당 질문에 답하기 위해, 정신질환의 유전화 현상을 이끌 수 있는 조직적 주체로서 정부와 기업을 가정하였다. 그리고, 정부와 기업 중 어떤 조직적 행위자가 정신질환의 유전화를 주도하였고, 시간에 따라 각 행위자의 영향력이 어떻게 변화하였는지 분석하였다.

이를 위해, 펍메드에서 정신질환을 다루는 논문을 수집하고 구조적 토픽모델링을 활용하여 정부와 기업이 정신질환의 의료 지식에 미치는 영향력을 분석하였다. 분석 결과를 요약하면 다음과 같다. 첫째, 정부는 뇌와 유전자로 대표되는 정신질환의 생물학적 원인을 파악하는 것에 집중한다. 둘째, 정부는 환자의 신체적 증상을 바탕으로 정신질환을 분류하고, 그들이 사회적 관계 안에서 겪는 고통에 주목한다. 셋째, 기업은 정신질환의 원인 및 증상이 아닌, 약물적 치료의 효과 및 안전성을 부각하는 것에 초점을 둔다. 넷째, 유전학 분야는 정부와 기업의 영향력이 동시에 작용하고 있다는 것을 보여준다. 출간 시기 변수를 통제했을 때, 유전학 연구는 정부의 펀드를 받은 논문에서 우세하게 나타난다. 하지만, 출간 시기와 기업 소속 변수의 상호작용 효과를 분석하면, 2010년대 들어 기업과 연계된 논문에서 유전학 연구의 비율이 증가하는 것을 확인할 수 있다.

본 연구의 분석 결과는 두 가지 함의를 갖는다. 첫째, 여타 분야에서 정부와 기업은 정신질환에 대한 의료 지식을 둘러싸고 일종의 역할 분담을 수행한다. 정부가 정신질환의 원인과 환자의 증상에 주목한다면, 기업은 약물 치료에 집중한다. 구체적으로, 정부는 정신질환의 원인을 파악하여 질병을 조기에 효과적으로 관리할 수 있다. 더불어, 환자들이 일상 생활에서 마주하는 신체적, 정신적 고통을 파악하여 환자들이 사회에 안정적으로 통합될 수 있도록 한다. 반면에, 기업은 약물의 안전성과 효과를 과학적으로 검증하여 의료 시장을 확대하는 것이 목표다. 정리하면, 정부는 의료 복지, 기업은 의료 시장의

확대를 목적으로 의료 지식을 구성한다.

둘째, 유전학은 정부와 기업의 이해관계가 결합할 수 있는 분야이다. 정부는 외현적인 증상이 아직 보이지 않더라도 유전적 위험 인자에 근거하여 환자 집단을 구성함으로써 정신질환을 효과적으로 관리할 수 있다. 기업은 개인의 유전적 특성에 근거하여 부작용을 줄이고 유전자의 작용을 조기에 억제할 수 있는 맞춤형 약물을 개발하여 약물의 소비층을 확대할 수 있다. 유전학이 정신의학, 약물학, 뇌-생물학 등 다양한 학문 분야에 걸쳐 정신질환 연구 내에서 광범위한 지식을 축적할 수 있었던 이유는 유전자에 대한 연구를 통해 정부와 기업이 동시에 이해관계를 충족할 수 있었기 때문이다.

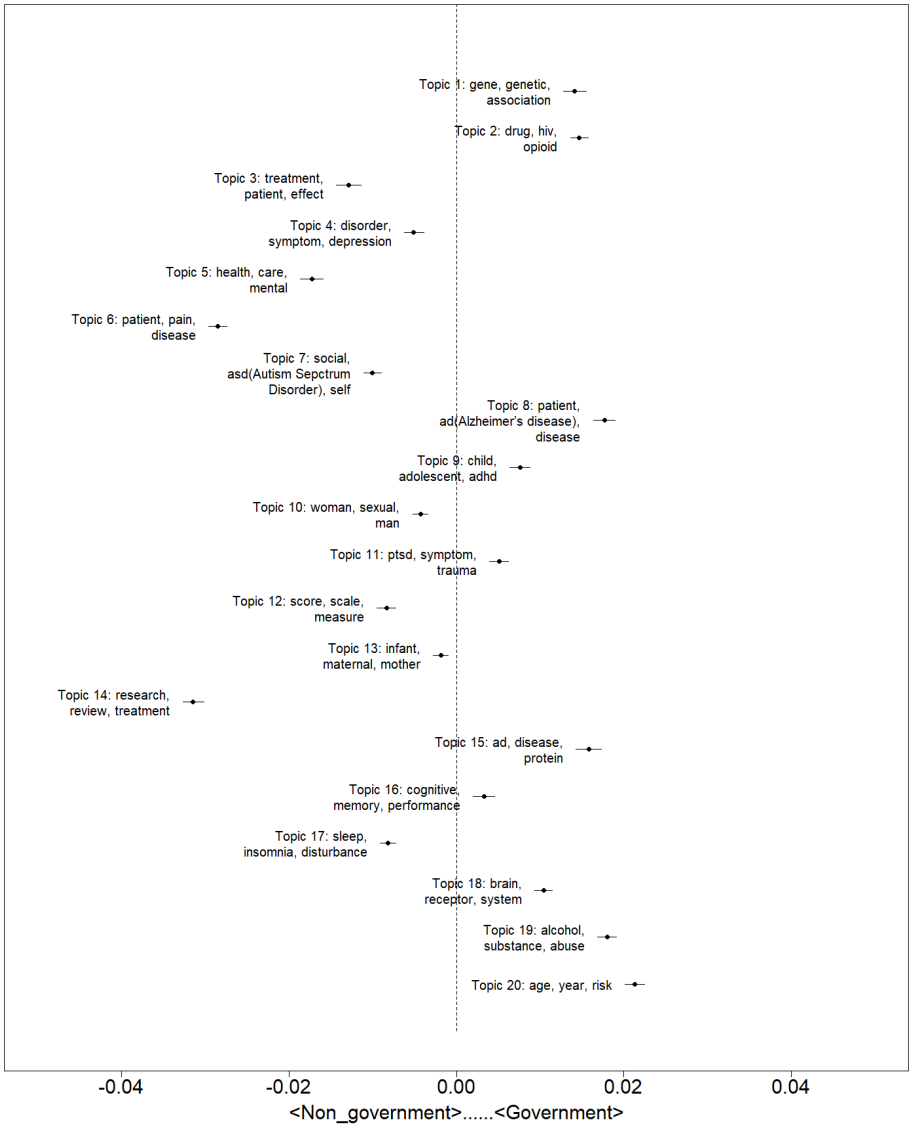
하지만, 본 연구는 두 가지 점에서 한계를 가진다. 첫째, 본 연구는 구체적으로 정부와 기업의 조직적 자본이 어떻게 결합하여 유전학 연구의 발전을 촉진하였는지 분석하지 못하였다. 이에 대해서는 정부의 후원을 받은 논문을 쓴 연구자들과 기업에 소속된 연구자들이 어떻게 네트워크를 구축하여 2010년대 이후의 변화를 이끌었는지 분석할 수 있을 것이다. 둘째, 본 연구는 유전학 연구를 둘러싼 정부와 기업의 영향력의 변화를 측정할 수 있었지만, 그러한 변화가 구체적으로 유전학의 연구 대상 및 내용에 어떤 변화를 일으켰는지 분석하지 못하였다. 이를 위해서는 시간의 변화에 따라 텍스트의 언어 구조가 어떻게 변화하는지 분석하는 동적 토픽모델링(Dynamic topic model)이나 동적 워드임베딩(Dynamic word embedding) 작업을 활용해볼 수 있다.

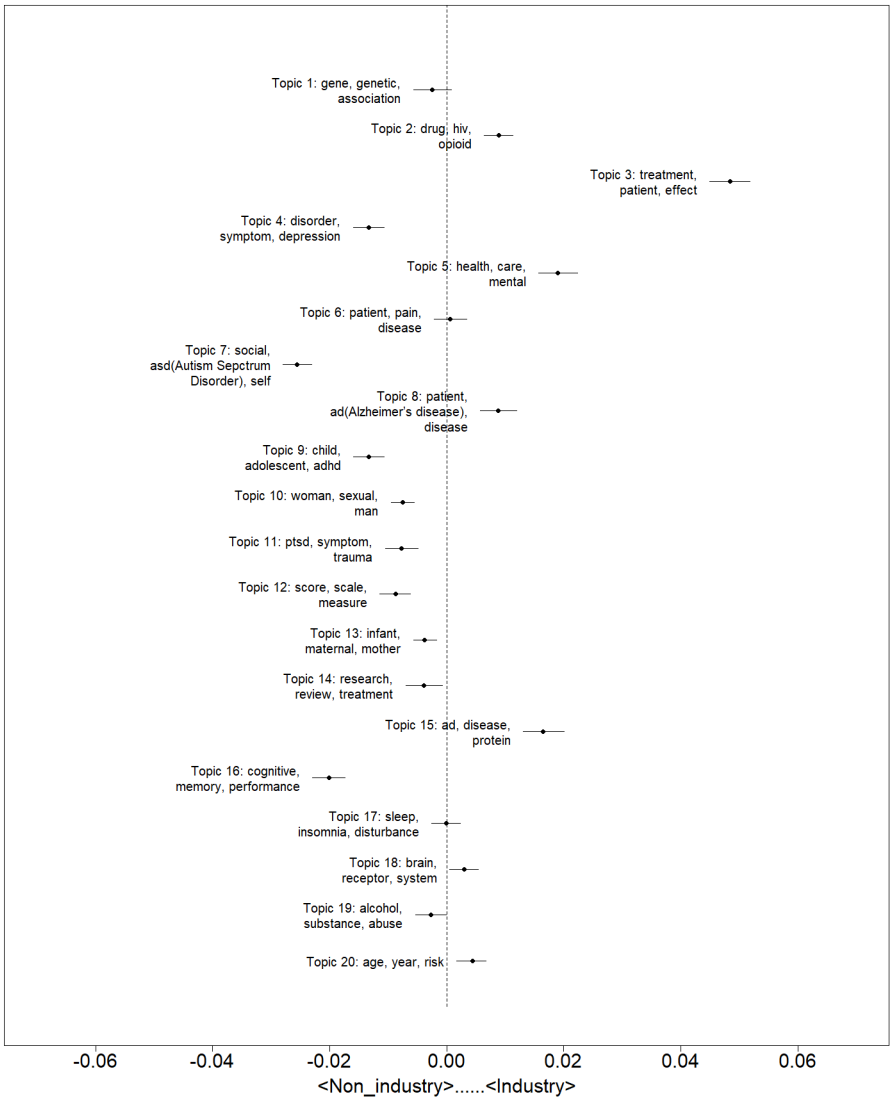
제 6 장 부록

부록 1 기업/비기업 변수 생성에 활용된 제약 및 바이오 기술 회사 목록

Roche	Mylan
Novartis	Teva Pharmaceutical Industries
Pfizer	CSL
Merck & Co.	Daiichi Sankyo
Bristol Myers Squibb	Otsuka Holdings
Johnson & Johnson	Meiji Holdings
Sanofi	Bausch Health Companies
AbbVie	UCB
GlaxoSmithKline	Alexion Pharmaceuticals
Takeda	Regeneron Pharmaceuticals
AstraZeneca	Eisai
Amgen	Servier
Gilead Sciences	Abbott
Eli Lilly	Grifols
Bayer	Fresenius Kabi
Novo Nordisk	Vertex Pharmaceuticals
Boehringer Ingelhiem	Chugai Pharmaceuticals
Allergan	Sumitomo Dainippon Pharma
Astellas Pharma	Ipsen
Biogen	Mitsubishi Tanabe Pharma

부록 2 정부/비정부 변수와 토픽 발현 확률의 관계(위) / 기업/비기업 변수와 토픽 발현 확률의 관계 (아래)




Topic 1: gene, genetic, association
Topic 2: drug, hiv, opioid
Topic 3: treatment, patient, effect
Topic 4: disorder, symptom, depression
Topic 5: health, care, mental
Topic 6: patient, pain, disease
Topic 7: social, asd(Autism Sepctrum Disorder), self
Topic 8: patient, ad(Alzheimer's disease), disease
Topic 9: child, adolescent, adhd
Topic 10: woman, sexual, man
Topic 11: ptsd, symptom, trauma
Topic 12: score, scale, measure
Topic 13: infant, maternal, mother
Topic 14: research, review, treatment
Topic 15: ad, disease, protein
Topic 16: cognitive, memory, performance
Topic 17: sleep, insomnia, disturbance
Topic 18: brain, receptor, system
Topic 19: alcohol, substance, abuse
Topic 20: age, year, risk
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06
<Non_industry>.....<Industry>

부록 3 토픽 별 주요 단어들

Topic 1	Highest Prob.	gene, genetic, association, study, mutation, syndrome, family, analysis, allele, schizophrenia
	FREX	rs, polymorphism, chromosome, snp, genome, haplotype, drd, loci, novo, duplication
Topic 2	Highest Prob.	drug, hiv, opioid, cocaine, infection, user, risk, among, prescription, high
	FREX	hiv, opioid, hcv, methadone, infect, aids, heroin, hepatitis, idu, immunodeficiency
Topic 3	Highest Prob.	treatment, patient, effect, study, medication, group, trial, week, therapy, dose
	FREX	placebo, ect, clozapine, risperidone, olanzapine, fluoxetine, discontinuation, mg, haloperidol, quetiapine
Topic 4	Highest Prob.	disorder, symptom, depression, patient, anxiety, psychiatric, diagnosis, depressive, major, criterion
	FREX	panic, ocd, obsessive, dsm, tic, psychosis, compulsive, mdd, psychotic, migraine
Topic 5	Highest Prob.	health, care, mental, service, program, intervention, treatment, need, medical, community
	FREX	staff, provider, nurse, service, cost, nursing, stigma, medicaid, professional, care
Topic 6	Highest Prob.	patient, pain, disease, case, cancer, year, delirium, chronic, mortality, hospital
	FREX	delirium, postoperative, surgical, icu, coronary, preoperative, cancer, myocardial, postoperatively, prostatectomy
Topic 7	Highest Prob.	social, asd, self, autism, participant, behavior, individual, negative, emotional, study
	FREX	emotion, empathy, attachment, asd, fear, reinforcer, music, attribution, rumination, mindfulness

Topic 8	Highest Prob.	patient, ad, disease, brain, alzheimer, subject, dementia, control, region, volume
	FREX	matter, pet, gray, fa, dlb, voxel, mci, volume, mri, spect
Topic 9	Highest Prob.	child, adolescent, adhd, behavior, problem, parent, age, family, school, year
	FREX	adhd, teacher, externalize, parenting, inattention, adolescence, parent, mph, oppositional, parental
Topic 10	Highest Prob.	woman, sexual, man, eat, female, weight, body, male, sex, control
	FREX	eat, bpd, nervosa, anorexia, bmi, eating, bulimia, bn, obese, overweight
Topic 11	Highest Prob.	ptsd, stress, symptom, trauma, experience, suicide, psychological, report, study, among
	FREX	ptsd, posttraumatic, violence, victimization, war, disaster, ipv, traumatic, suicide, deployment
Topic 12	Highest Prob.	score, scale, measure, factor, test, assessment, sample, study, analysis, item
	FREX	reliability, qol, validity, psychometric, bdi, ret, instrument, hrqol, item, mmpi
Topic 13	Highest Prob.	infant, maternal, mother, exposure, pregnancy, birth, woman, developmental, delay, month
	FREX	infant, pregnancy, birth, postpartum, prenatal, neonatal, preterm, perinatal, gestational, newborn
Topic 14	Highest Prob.	research, review, treatment, study, clinical, evidence, approach, article, include, base
	FREX	article, literature, paper, scientific, science, methodological, framework, approach, search, systematic
Topic 15	Highest Prob.	ad, disease, protein, beta, abeta, cell, level, amyloid, alzheimer, tau
	FREX	abeta, app, prion, secretase, phosphorylation, prp, soluble, deposit, fibril, fatty
Topic 16	Highest Prob.	cognitive, memory, performance, task, test, control, deficit, group, function, impairment

	FREX	phonological, erp, auditory, executive, dyslexic, fmri, stroop, dmn, word, fluency
Topic 17	Highest Prob.	sleep, insomnia, disturbance, time, quality, movement, study, night, hour, subject
	FREX	sleep, night, sleepiness, daytime, rls, deprivation, nocturnal, circadian, wake, rem
Topic 18	Highest Prob.	brain, receptor, system, mechanism, human, effect, function, hd, induce, role
	FREX	hd, neurotransmitter, gaba, da, pituitary, glutamate, hpa, nmda, adrenal, glutamatergic
Topic 19	Highest Prob.	alcohol, substance, abuse, dependence, drug, problem, smoking, student, drinking, among
	FREX	alcohol, drinking, gambling, tobacco, drink, smoker, smoke, cannabis, cigarette, sud
Topic 20	Highest Prob.	age, year, risk, study, ci, associate, cognitive, old, high, participant
	FREX	ci, adjust, confidence, decline, interval, ratio, odd, old, regression, mmse

부록 4 토픽의 제목

토픽 순서	토픽 제목
1	인간 유전자와 정신질환의 연관성
2	약물 사용자의 감염성 질환
3	정신질환 약물의 효과와 안전성
4	정신질환 환자의 신체적 증상
5	공동체 치료 및 간호
6	수술 후 합병증
7	감정 및 사회적 행동 장애
8	치매(알츠하이머)와 뇌 사진술
9	아동과 청소년의 행동 장애 (ADHD)
10	섭식장애
11	외상 후 스트레스 장애
12	성격 측정의 신뢰성
13	저체중 출생아의 신경발달장애
14	심리치료
15	인간 뇌에서 치매의 병리학적 특성
16	언어 장애
17	수면 장애
18	신경정신병의 병리학적 특성
19	음주, 흡연, 약물 중독
20	노인 인구에서 신체 활동과 인지 기능의 연관성

참고문헌

Busfield, J. (2006). Pills, power, people: sociological understandings of the pharmaceutical industry. *Sociology, 40*(2), 297-314.

Can genes be patented?: MedlinePlus Genetics. (n.d.). MedlinePlus. Retrieved August 7, 2022, from https://medlineplus.gov/genetics/understanding/testing/genepatents/

Chang, K.-L., Weitzel, K., & Schmidt, S. (2015). Pharmacogenetics: using genetic information to guide drug therapy. *American family physician, 92*(7), 588.

Clarke, A. E., Shim, J. K., Mamo, L., Fosket, J. R., & Fishman, J. R. (2010). 1/Biomedicalization Technoscientific Transformations of Health, Illness, and US Biomedicine. *Biomedicalization*, 47-87.

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., . . . MacArthur, D. G. (2020). A brief history of human disease genetics. *Nature, 577*(7789), 179-189.

Conrad, P. (2005). The Shifting Engines of Medicalization. *Journal of Health and Social Behavior, 46*(1), 3-14. doi:10.1177/002214650504600102

Cook-Deegan, R., & Heaney, C. (2010). Patents in genomics and human genetics. *Annual review of genomics and human genetics, 11*, 383.

Evans, J. A. (2010). Industry induces academic science to know less about more. *American Journal of Sociology, 116*(2), 389-452.

Frances, A. (2014). *정신병을 만드는 사람들 : 한 정신 의학자의 정신병 산업에 대한 경고*. 서울: 서울 : 사이언스북스.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis, 21*(3), 267-297.

Guo, G., Roettger, M. E., & Cai, T. (2008). The integration of genetic propensities into social-control models of delinquency and violence among male youths. *American Sociological Review, 73*(4), 543-568.

Heller, M. A., & Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science, 280*(5364), 698-701.

Hyman, S. E. (2000). The genetics of mental illness: implications for practice. *Bulletin of the World Health Organization, 78*, 455-463.

Hyman, S. E. (2010). The diagnosis of mental disorders: the problem

of reification. *Annual review of clinical psychology, 6*.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R* (Vol. 103). New York: New York: Springer.
Kuhn, T. S. (1986). *과학혁명의 구조*. 서울: 서울 : 梨花女子大學校 出版部.
Latour, B. (2016). *젊은 과학의 전선 : 테크노사이언스와 행위자 - 연결망의 구축*. 파주: 파주 : 아카넷.
Lee, T. J. (1993). A medico-historical review on biomedicine and its limitations and problems. *Korean Journal of Medical History, 2*(1), 1-9.
Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. The Lancet, 373(9659), 234-239.
Masic, I., & Milinovic, K. (2012). On-line BiOmedical dataBases-the Best sOurce fOr Quick search Of the scientific infOrmatiOn in the BiOmedicine. *Acta Informatica Medica, 20*(2), 72.
Nunnally, A. C., Webster, C. J., Brown, S. A., & Cohen, G. A. (2005). Genetic patent protection in the pharmaceutical and biotechnology industries. *Public Health Genomics, 8*(4), 209-216.
Pezawas, L., Meyer-Lindenberg, A., Drabant, E. M., Verchinski, B. A., Munoz, K. E., Kolachana, B. S., . . . Weinberger, D. R. (2005). 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nature Neuroscience, 8*(6), 828-834.
PubMed Overview. (n.d.). PubMed. Retrieved August 7, 2022, from https://pubmed.ncbi.nlm.nih.gov/about/
Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association, 111*(515), 988-1003.
Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software, 91*(1), 1-40.
Rule, A. (2018). *Standpoints on Psychiatric Deinstitutionalization*: Columbia University.
Sanders, A. R., Duan, J., Levinson, D. F., Shi, J., He, D., Hou, C., . . . Olincy, A. (2008). No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *American Journal of Psychiatry, 165*(4), 497-506.

Shorter, E. (2020). *정신의학의 역사 : 광인의 수용소에서 프로작의 시대까지* (개정판. ed.). 서울: 서울 : 바다출판사.

The National Institute of Mental Health Strategic Plan. (2022, June). National Institute of Mental Health (NIMH). Retrieved August 7, 2022, from https://www.nimh.nih.gov/about/strategic-planning-reports

김란우, & 송수연. (2020). 한국 학계의 고유성은 존재하는가? 한국 사회학과 국제 사회학의 지식 담론 구조 비교를 중심으로. *한국사회학, 54*(4), 1-40.

송준모, & 강정한. (2018). 메갈리아의 두 딸들. [Two Daughters of Megalia : Identity Differentiation in Response to the Anonymity Level in Online Community]. *한국사회학, 52*(4), 161-206. doi:10.21562/kjs.2018.11.52.4.161

심창구, & Kyōto Daigaku. Daigakuin, Y. (2012). *새로운 약은 어떻게 창조되나*. 서울: 서울 : 서울대학교 출판문화원.

양병환. (2017). *정신병리학 특강*. 파주: 파주 : 집문당.

조병희. (2006). *질병과 의료의 사회학*. 파주: 파주 : 집문당.

조원광. (2017). 의료 지식의 사회적 구성과 건강 행동 : 유방암 논문 초록과 환자 포럼 게시글의 토픽 모델링을 중심으로. Retrieved from http://snu-primo.hosted.exlibrisgroup.com/82SNU:82SNU_SSPACE210371/136987

Abstract

Who Leads the Genetification of Mental Disorders

: Focusing on the role of Government and Medical Industry

Jungsik Choi
Department of Sociology
The Graduate School
Seoul National University

This study defines the genetification of mental disorders as a phenomenon in which genetics and psychiatry are combined to understand and explain mental disorders in terms of genes and their interactions. How was genetic research able to accumulate comprehensive medical evidence for the diagnosis, prediction, and treatments of mental disorders? Who are the organizational actors that produce the scientific knowledge on genetic research by mobilizing various human, physical, and technological resources?

In order to answer these questions, this study analyzed the impacts of governmental agencies and medical industries

(pharmaceutical and biomedical companies) on the medical knowledge of mental disorders. After collecting the articles dealing with mental disorders from PubMed and applying structural topic modeling to the textual data, this study analyzed how publication dates, governmental funding, and affiliation with the medical industry influenced the topic prevalence.

The statistical findings are as follows. First, the governmental agencies focused on the two biological causes – brain and genes – of mental disorders. Second, they organized patient groups based on their physical symptoms and pay attention to the pain they face in their social relationships. Third, the medical industries focus on validating the efficacy and safety of the drugs for mental disorders. Fourth, when the publication date variable is controlled, the governmental agencies led the development of genetic research. However, when the interaction effect of the publication dates and governmental funding variables is observed, genetic research has increased its proportion in the articles affiliated with the medical industry since 2010.

The result has two implications. First, the governmental agencies and the medical industry played different roles in the medical knowledge of mental disorders. Whereas the governmental agencies focused on the causes and the symptoms of mental disorders, the medical industry paid attention to the safety and efficacy of the drug treatments. Second, governmental and industrial interests are combined in the field of genetics. The governmental agencies can manage mental disorders effectively based on the genetic risk elements. The medical industry can expand the medical market by developing customized drugs that can reduce their side effects and inhibit the genetic mechanisms early.

Keywords : Mental disorders, Genetification, Structural topic modelling, Government, Industry
Student Number : 2018-28122