

인공지능과 로봇의사윤리

맹주만*

주제분류 윤리학, 응용윤리학, 형이상학

주요어 인공지능, 로봇의사, 로봇의사윤리, 의료윤리, 의사윤리, 자율성과 상호 동의의 원칙, 악행 금지와 공동 책임의 원칙, 선행과 돌봄의 원칙, 의료권과 정의의 원칙

요약문

현행 의사의 직업윤리는 의사의 윤리적 소명과 책무 및 의사와 환자 사이의 윤리적 관계 중심의 원칙을 담고 있다. 가까운 장래에 의료인공지능이 탑재된 로봇의사는 인간의사의 질병 진단과 처방에 중요한 역할을 할 것이다. 이와 함께 인간의사의 로봇의사에 대한 의존도가 커질수록 더욱 더 많은 윤리적 문제들이 발생할 것이다. 그럴 경우에 이러한 문제들을 해결하는데 인간의사 중심의 의료윤리만으로는 한계가 있다. 이에 나는 로봇의사윤리의 제정이 필요하다고 생각한다. 의료행위에 인공지능의 도움이 불가피한 시대에서 예상되는 인간의사와의 협업 관계를 고려하면 더욱 그렇다. 이 글은 이를 위한 예비적 고찰이다. 나는 미래의 로봇의사에게 적용되어야 할 가능한 의료윤리의 근본 원칙을 로봇의사의 윤리원칙이라는 이름 아래 제시할 것이다.

* 중앙대학교

1. 인공지능, 로봇의사, 의료윤리

나는 로봇의사를 인간의 생명과 건강을 증진시키는 의료적 목적을 잘 수행하도록 제작된 의료인공지능 기계로 정의한다. 그리고 나는 지금은 데스크 탑 컴퓨터처럼 특정한 위치에 놓여 있는 불박이 기계에 불과지만, 앞으로 이 의료기계는 안드로이드(android) 로봇의 형태로 제작되어 많은 면에서 인간의사와 유사한 방식으로 의료활동을 하게 되리라 가정한다. 이러한 목적 하에 제작된 로봇의사는 최소한 한 가지 분야 이상에서 의료적 진단과 치료법의 제시를 통해서 인간의사의 역할을 보조 및 대행하게 될 것이다. 실제로 이런 로봇의사를 만나러 갈 때가 멀지 않았다. 아니 이미 시작되었다. 머지않아 머신러닝 혹은 딥러닝과 빅데이터 기술이 결합된 인공지능을 장착한 로봇의사가 미래의료를 선도할 것이다.

4차 산업혁명의 급속한 성장은 이제 로봇의사의 진화 속도 또한 가능하기 어렵게 만들고 있다. 바둑에 특화된(specialized) 알파고와 같이 한 가지 기능과 목적에 맞게 프로그래밍 된 특수인공지능(ANI; Artificial narrow Intelligence) 혹은 약인공지능, 그리고 2030년쯤이면 실현될 것으로 예견되는 인간 수준에 버금가는 인지 능력과 기능을 갖춘 강인공지능 혹은 범용인공지능(AGI; Artificial general Intelligence)은 물론 이보다 1,000배 이상으로 고도화된 초지능(superintelligence) 혹은 초인공지능(ASI; Artificial super Intelligence)의 도래는 의료 보건 산업에 엄청난 변화를 가져올 것이다.¹⁾ 심지어 진료와 진단 및 처치까지 하는 로봇외과의사, CEO가 직원을 고용해서 회사를 운영하듯 로봇의사들이 진료하는

1) 초지능 및 인공지능의 등급에 대해서는 N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, 26, 63-64쪽; 박영숙·제롬 글렌, 『세계미래보고서 2055』, 114쪽; 김재인, 『인공지능의 시대, 인간을 다시 묻는다』, 9쪽; 최윤섭, 『의료 인공지능』, 47-56쪽; ‘artificial intelligence’ in wikipedia 참조.

병원도 등장할 것이다. 무엇보다도 환자 맞춤형 진단과 치료 및 신약 개발은 인간의 생명과 수명의 연장 등 인간의학의 미래를 완전히 바꾸어 놓을 것이다.

로봇의사의 등장과 함께 미래의학과 의료보건산업에서 예측되는 변화는 그냥 있을 법한 가상 시나리오가 아니다. 이미 우리는 그 도상에 서 있다. 실제로 2012년 IBM에서 개발한 슈퍼컴퓨터 의사 왓슨(Doctor Watson)의 등장으로 그 서막을 알린 바 있다. 이미 한국의 가천 길병원에서도 2016년 12월 IBM의 왓슨을 도입해 운영하고 있으며, 다른 국내 병원들의 왓슨 도입도 이어지고 있다. 2016년 2014년 엠디앤더슨 암센터의 발표에 따르면 암 전문의 왓슨의 폐암 진단 및 치료방법 선택의 정확도는 82.3%에 이른다. 2015년 기준 한국인의 암 사망 중 1위를 차지한 폐암의 오진율은 19%에 달한다. 이렇게 보면, 로봇과 로봇, 로봇과 인간 사이의 협력이 더 한층 강화되어 그야말로 막대한 양의 정보를 갖게 될 ‘협력형 로봇’으로서 로봇의사의 검진 정확도는 갈수록 더 높아질 것이다.²⁾

최근에 『특이점이 온다』(*The Singularity is Near*, 2017)는 저서의 출간으로 세계적인 주목을 받은 발명가이자 미래학자인 레이 커즈와일(Ray Kurzweil)은 이 저서를 시작하면서 27년 전 『지적 기계의 시대』(1990)의 마지막 부분에서 그 자신이 이미 “나는 21세기 전반에 인간 지능과 구별되지 않는 수준의 기계 지능이 등장할 것이라 전망했다”³⁾는 것을 회고하고 있다. 이렇게 그는 새로운 혁명이 시작된 지 매우 오래되었음을 새삼 강조하고 있다. 우리는 오늘날 그 혁명을 제4차 산업혁명이라고 부른다. 이를 이미 하나의 역사로서 정리해 놓은 인터넷 백과사전 ‘위키피디아’에 따르면, 제4차 산업혁명은 정보통신기술(ICT)의 융합으로 이루어낸 혁명 시대를 말한다. 그리고 이 혁명을 주도하고 있는 핵심 기술이 인공지능

2) 박영숙·제롬 글렌, 『세계미래보고서 2055』, 43-44쪽.

3) 레이 커즈와일, 『특이점이 온다』, 18쪽.

이다. 이 인공지능 기술은 빅 데이터 분석, 로봇공학, 사물인터넷, 무인운송수단(무인 항공기, 무인 자동차), 3차원 인쇄, 나노 기술 등과 같은 분야에서 새로운 기술 혁신들을 주도하고 있다. 이는 디지털, 물리학, 생물학 등 경계가 없어지고 융합되는 기술 혁명으로서 앞으로 더 가속화될 것이다. 구체적으로는 인체의 정보를 디지털 세계에 접목하는 기술인 스마트워치나 스마트 밴드를 이용한 모바일 헬스케어, 로봇 공학, 인공지능, 나노 기술 등은 수십억 명의 사람들을 계속해서 웹에 연결하고 비즈니스 및 조직의 효율성을 획기적으로 향상시키며 더 나은 자산 관리를 통해 자연 환경을 재생산할 수 있는 커다란 잠재력을 가지고 있는 것으로 평가된다.

주지하듯이 1차 산업혁명은 철강 산업과 증기기관의 개발로 가능해진 기계생산, 2차 산업혁명은 전기 기술의 발명으로 기계생산 산업이 사회 전반으로 확장된 대량생산, 3차는 디지털 혁명으로 불리는 컴퓨터와 인터넷의 등장으로 자동생산이 실현되었다면, 4차 산업혁명의 키워드는 바로 인공지능이 구현하는 자율생산이다. 자율생산의 사례는 이미 사물인터넷과 같은 표현을 통해서 익숙해졌다. 현재 수준에서도 기계가 갖추고 있는 이러한 자율성은 인공지능을 전후 상황에 대한 자율적 판단 기능까지 겸비한 인공지능 즉 상황지능(contextual intelligence)의 수준까지 도달했다. 이는 단순 인공지능 수준을 훌쩍 뛰어 넘는 AGI 수준의 로봇의 출현이 가능하다는 것을 시사한다. 이런 변화의 추이는 특히 인간의 건강을 책임지는 의료분야에서 더욱 위력을 발휘하게 될 것이다. 1~2년에 한 번씩 하는 정기적인 종합검진을 상황지능 로봇의사는 매일 매일 대신 해줄 수 있으며, 더욱이 사람들과 함께 생활하면서 수시로 건강을 체크하고 예방하는 비서형 내지는 가족형 로봇의사를 편의시설 갖추듯 구비할 수도 있게 될 것이다. 유전자, 유전 형질 등을 기록한 보건 데이터의 이용은 오류의 가능성을 최소화한 맞춤형 치료를 가능하게 할 것이다.

상황판단과 의사소통을 겸비한 실질적 자율형 AGI 로봇의사는 많은

의료 부문에서 집적된 의료 지식과 기술을 이용해서 인간의사보다 의료 행위를 더 잘 해내게 될 것이다. 거부할 수 없는 제4차 산업혁명의 파고가 사회 전반에 걸쳐 엄청난 혁신을 가져오고 있으며, 인간의 삶의 조건과 지형을 급속도로 바꾸어 놓고 있다. 그리고 다른 분야에 비하면 인간의 생명과 안전을 다루기 때문에 적응과 수용이 조심스럽고 늦을 수밖에 없는 의료 보건 분야가 최종 귀착지가 될 것이다. 하지만 예상할 수 있듯이 그와 함께 의료 행위와 관련된 많은 윤리적 문제들도 속출할 공산이 크다. 그 중에서도 특히 원론적으로 말하자면 AGI 수준만이 아니라 ASI 수준으로까지 진화하게 될 자율형 로봇의사는 인간 내지는 인간의사의 통제를 벗어날 가능성도 안고 있다. 인간 생명을 다룬다는 점에서 이러한 가능성은 가장 경계해야 할 의료윤리적 문제를 초래할 것이며, 이에 대한 심각한 고민도 이미 시작되었다.

2. 의료인공지능과 증거기반의학

인공지능의 진화는 어디까지 이루어질 것인가? 인간 수준의 AGI의 등장으로 예견되는 놀라운 변화들에는 기계가 인간을 위협하게 되는 시나리오들이 적지 않다. 이처럼 인류의 재앙 같은 파괴적인 예측이 아니더라도 일견 건설적인 예측이면서도 두려움을 갖게 하는 마음에는 인간의 능력을 대신하거나 모든 면에서 인간보다 더 잘 할 수 인공지능의 등장이 자리 잡고 있다. 이 중에는 인간의 고유성을 강조하는 희망적인 예견도 많다. 일례로 마윈 알리바바 회장은 2016년 3월 중국 경제발전 고위급포럼에서 “기계가 인류보다 강해지고 똑똑해질 것이라고 생각한다. 그러나 기계가 인류보다 절대로 지혜로울 수는 없다. ... 지혜와 정신, 마음은 인간만의 특성이며, 컴퓨터는 성취감과 우정, 사랑 같은 감정을 느낄 수 없다.”⁴⁾ 하는 등 인공지능의 발달에도 불구하고 인간 고유의 영역

을 강조하기도 한다. 그러나 다양하면서도 엇갈리는 진단과 평가들이 어 떠하던 개별 인간의 특수한 능력을 뛰어넘는 인공지능의 등장은 필연적 이다.

인간의 직능을 대체할 수 있는 인공지능의 등장으로 일어날 수 있는 혁신적인 변화를 가장 잘 실감할 수 있는 부문이 의료분야의 인공지능이 다. 비노드 코슬라 썬마이크로시스템즈의 공동 창업자는 “기술의 발전으 로 앞으로 대부분의 의사가 로봇의사(Doctor Algorithm 또는 Dr. A)에게 자리를 내줘야 할 것이며,”⁵⁾ “앞으로 테크놀로지가 의사의 80%를 대체 할 것이다.”⁶⁾라는 충격적인 진단을 내놓기도 했다. 이는 암진단 뿐만 아 니라 인공지능 기술발전과 함께 배아 대상 유전자가위기술 적용 등 유전 자치료연구, 세 부모아이, 인공배아생성 등 생식세포를 이용한 새로운 융 합 연구, 이종장기이식, 키메라연구, 안면이식 등 새로운 장기이식 연구, 빅 데이터 기술을 포함한 데이터 과학 및 디지털헬스의 발전 가속화는 피할 수 없으며, 이와 함께 개인 종합 정보의 의료 전산화 및 유출 가능 성도 심화될 것이다. 이런 변화 가운데 상당 부분은 이미 실현되었거나 과거의 일이 되버렸다.⁷⁾ 그 가운데 이와 같은 일련의 변화를 상징적으로 보여주는 것이 로봇의사 왓슨의 출현이라 할 수 있다.

2013년에 발표된 한 기사에 의하면, “의사의 잘못된 진단으로 목숨을 잃는 사람이 한 해에 미국에서만 40,500명이 넘는다는 연구가 있다. 이 수치는 유방암으로 인한 사망률과 비슷하다.”⁸⁾ 이 기사는 또한 미국의 의학연구소는 2000년 발표한 ‘인간이기에 실수를 한다’라는 제목의 보고 서를 통해 환자 안전의 중요성을 강조했고, 2011년 발표한 보고서 ‘헬스

4) http://biz.chosun.com/site/data/html_dir/2016/03/20/2016032000637.html

5) <https://techcrunch.com/2012/01/10/doctors-or-algorithms>

6) <https://www.youtube.com/watch?v=TTcCkzrjNLY&feature>

7) 이에 대해서는 J. Rifkin, J., *The Biotech Century*, 1998; Richard W. Oliver, *The Coming Biotech Age*, 2000.

8) <http://platum.kr/archives/6985>

IT와 환자 안전'에서 기술이 매우 중요한 역할을 할 것이라 밝혔다고 알려준다. 빅데이터와 같은 방대한 의학적 자료들로부터 최적의 진단과 치료법을 찾아내는 로봇의사의 의료적 접근 방식이 증거기반의학과 유사하다. 더욱이 인간의사와 로봇의사의 협업은 단순한 메커니즘 추론이나 임상적 전문가의 경험적 판단에 비해 관찰과 실험에 근거한 비교임상연구와 더 잘 부합한다. 실제로 미래의 로봇의사는 증거기반의학의 종합적인 구현체가 될 수 있다.

환자에 관한 개인 정보, 기존치료방법과 임상결과 등이 총망라되어 있는 빅데이터 의료정보, 인간의사 개인이 접근하기 어려운 최신 의학지식과 정보의 지속적인 업데이트, 이에 기초한 최적의 치료법의 제시 등 대부분의 의료지식과 기술정보들이 증거기반의학의 목표와 정확히 일치한다. 이는 그동안 의료업계의 현장에 적용되던 경험중심의학이라 할 수 있는 “관행기반의학”(practice-based medicine)으로부터 과학적 실험과 관찰 결과 중심의 “증거기반의학”(evidence-based medicine)으로의 이행 및 전환과 맥을 같이 한다고 할 수 있다.⁹⁾

빅 데이터 기술을 이용한 인공지능 컴퓨터에 의한 진단의 현실화, 즉 빅데이터가 결합된 인공지능-로봇에 의한 의사결정, 환자 건강도우미 ‘헬스아바타’의 현실화와 같은 미래의료 환경의 변화와 선도는 전문 의사의 개념을 뒤바꿔놓을 것이며, 심지어는 현재 의사가 하는 일들 중 많은 부분을 의사 면허 없이도 대체할 수 있게 될지도 모른다. 아마 의학교육 및 의료전문가 양성제도가 완전히 바뀌게 될 것이다. 그렇게 되면 인간 의사와 로봇의사 사이에서 진료와 치료 업무의 분업화와 차별화가 가속될 수도 있다. 그리고 인간의사와 로봇의사 각자의 장점을 살린 전담 분야들이 특성화될 수도 있다. 가령, 이성과 감성으로 대별되는 로봇의사와 인간의사의 역할 분담 및 전문화가 그 좋은 예가 될 것이다. 인간의사와

9) 제러미 하워, 『증거기반의학의 철학 : 의학의 새로운 패러다임』, 36-57쪽 참조.

로봇의사가 지닐 수 있는 상대적 특성과 장점을 좀 더 구체적으로 비교해보면, 먼저 인간의사의 경우에 제한된 정보, 경험, 직관에의 의존성이 높기 때문에 환자 개인의 특수성에 대한 고려가 의사 마다 다를 수 있으며, 이와 함께 피로와 집중력에 따라서 부정확한 진단 또는 치료방법 선택 시 실수나 오류의 가능성이 로봇의사 보다 상대적으로 높을 수 있지만, 환자에 대한 인격적 대면, 예술, 윤리적 접근 및 대면 맞춤 진료와 상담은 더욱 강조될 수 있다. 반면에 로봇의사는 방대한 양의 정보, 과학적 데이터, 방대한 의료지식과 임상결과의 활용, 최적화된 치료법 제시, 일정한 확률로 정확한 진단 및 동일 질병에 동일한 처방 등 기술적인 측면에서 의학적 오류를 최소화할 수 있을 것이다.

3. 의료윤리와 로봇의사

인간지능과 인공지능을 구분할 때, 인공지능이 범접할 수 없는 인간 영역이 있다. 칸트가 말하듯이 인간 혹은 인간지능은 스스로 목적을 정립하는 자기 목적적 존재다. 인간적 의지는 “어떤 법칙의 표상에 따라 자기 자신의 행위를 규정하는 능력”¹⁰⁾으로서 상대적 가치만을 갖는 ‘사물’과는 다르다. 말하자면 인간지능은 무엇이 문제인지 스스로 결정하고 만들어내기도 하지만, 인공지능은 누군가가 문제가 무엇인지 알려주어야만 그 문제를 해결할 수 있다. 즉, 인간지능은 문제 환경을 능동적으로 지각하고 이해하면서 스스로 직면한 문제를 발견하고 해결하지만, 인공지능은 인간에 의해서 부과된 과제 환경 안에서 정해진 문제를 해결한다.¹¹⁾ 따라서 기본적으로 인간지능은 스스로 목표와 문제를 설정하지만,

10) I. Kant, *Grundlegung zur Metaphysik der Sitten*, 59쪽

11) I. Kant, *Grundlegung zur Metaphysik der Sitten*, 59쪽 이하; 김재인, 『인공지능의 시대, 인간을 다시 묻는다』, 69-78; 박충식, 「생명으로서의 인공지능: 정보철학적 관점에서」, 21-40쪽.

인공지능은 스스로 목표를 정할 수 없으며, 인간에 의해서 일정한 목표와 문제 내지는 과제가 먼저 부여되어야 한다. 그러나 의학적 목표로 제한되는 로봇의사의 경우에는 이런 문제가 불식될 수 있다. 인간의 생명과 건강이라는 일정한 목적을 위해 프로그래밍 된 로봇의사는 이러한 목적 아래서 행동할 것이기 때문이다. 반면에 일반적으로 진정한 의미의 감성로봇 혹은 감성로봇의 가능성은 낮다.¹²⁾ 그러나 지식과 기술의 측면에서 이런 특화된 목적지향적 로봇의사가 가능하다면 적어도 동일한 목적과 기능에 한정된 감성로봇, 즉 감성 로봇의사는 가능하다고 생각된다. 이 때의 감성로봇 내지는 감성로봇의사 역시 인간의 이성적 기능과 통합적으로 작동하는 감정능력의 수준에는 이를 수 없다고 해도 특수한 감정지각에 반응하는 특화된 로봇 내지는 로봇의사는 존재할 수 있으며, 이 역시 제한적이긴 하지만 감성로봇의사로서의 일정한 역할을 기대할 수 있을 것이다. 다시 말해 인간의 섬세하고 민감한 감정 인지에는 한계를 지니겠지만, 의학적 치료에 한정된 환자를 상대하는 목적지향적 맞춤형 감성로봇의사는 가능할 것이다.

이렇게 미래의 인공지능 역시 인간처럼 목적과 문제까지 스스로 설정하는 수준까지 진화할 가능성은 낮다고 하더라도, 따라서 최소한 인간기능과 인공지능 사이에 결코 뛰어 넘을 수 없는 간극이 존재한다는 것을 인정하더라도, 현재의 급속한 인공지능 기술의 발전을 감안할 때 강인공지능 이상으로 특정한 목적과 목표를 수행하는데 적합한 이른바 특화된 초인공지능으로의 진화는 가능하다고 보여진다. 이러한 수준의 로봇을 나는 기능형 혹은 목적형 로봇이라 부른다. 아마 로봇의사 역시 이러한 유형의 인공지능로봇에 해당될 것이다. 다시 말해서 인공지능 기술의 발달이 궤도 이르게 되면 특수한 질환 혹은 몇몇 연계질환의 진단과 치료에 특화된 강인공지능(범용 혹은 일반 인공지능) 내지는 초인공지능 수준의 로

12) 천현득, 「인공지능은 감정을 가질 수 있을까?」, 137-176쪽.

봇의사가 인간의 생명과 건강이라는 일반적 목표에 부합하는 진단과 진료 및 처방을 주도할 것이다. 이런 시대가 도래하면 예상치 못한 많은 윤리적 문제들이 생길 수 있으며, 더불어 인간 의사와 로봇 의사 양자의 윤리적 관계도 중요한 문제로 부각할 것이다.

기능형 및 목적형 로봇 의사는 생명과 건강 증진이라는 특수한 목적에 알맞도록 프로그램 된 인공지능이다. 장차 이런 로봇 의사는 외형상 인간의 모습을 한 멋진 로봇으로 발전할 것이며, 현재 삼성 갤럭시 스마트폰으로 음성 대화를 나누는 방식처럼 상호 소통이 가능한 모습까지 갖추게 될 것이다. 이러한 수준의 로봇 의사 단독의 진료가 어디까지 가능할지 예단하기는 어렵다. 다만 관건은 이런 수준의 로봇 의사가 목적 일반의 설정까지는 불가능하더라도 의료 행위와 같은 사전에 설정된 특수한 목적을 실천하는 수준의 인공지능이 가능하다면 이런 행위형 로봇 의사에게는 모종의 책임과 권리를 행위 주체로서의 지위가 부여될 수도 있을 것이다.¹³⁾ 더불어 인간 의사는 특화된 목적에 맞춰 제작된 로봇 의사와 협업을 하거나 로봇 의사의 도움에 절대적으로 의존하게 될 것이며, 로봇 의사는 자연스럽게 의사 공동체의 일원이 되어 있을 것이다. 이렇게 되면 로봇 의사 역시 인간 의사가 져야 할 윤리적 문제들로부터 자유롭지 못하게 될 것이며, 또 그렇게 되어야 한다.

13) 인공지능 프로그램이나 지능형 로봇으로서 로봇 의사가 행위의 주체로 인정받으려면 외부 대상에 대한 지각과 그것에 관한 동작의 협응, 즉 “감수자 체계와 행위자 체계의 통합”이 필수적일 뿐만 아니라 이러한 협응이나 통합의 배후에는 행위를 낳는 의도를 지닌 주체적 존재, 즉 자아의 존재 또한 필수적이다. 그러나 나는 초인공지능으로서 의료의 목적에 특화된 로봇 의사는 현재와 같이 문서로 주고받는 상호 관계를 넘어서 언어적 소통을 통한 인간 의사와의 상호 관계가 가능하며, 따라서 최소한 특수한 목적에 맞게 제작된 인공지능 로봇 의사는 언어적 주체로서 간주될 수 있다고 생각한다. 그리고 만일 로봇 의사가 제한적으로 언어적 주체로 간주될 수 있다면 이는 동시에 그 존재가 행위 주체로 간주될 수 있다는 것을 함축한다. 인공지능 일반과 관련하여 이러한 가능성에 대해서 “부정적이거나 기껏해야 유보적인” 견해에 대해서는 다음을 참조. 고인석, 「인공지능의 존재 지위에 대한 두 물음」, 161-181쪽.

모든 면에 인간과 동일한 수준의 윤리로봇이 아닌 목적형 로봇의사일 지라도 진단과 처치의 주체 및 그에 따른 책임 문제 등이 불거질 것이다. 현실적으로는 인간의사가 의료적 행위의 주체이자 최종 결정권자이며 로봇의사는 선택적인 협업 대상으로서 보조 역할 내지는 자문 역할에 한정될 공산이 크다. 그러나 이런 인위적인 관계 설정이 생각대로 유지될 가능성은 희박해 보인다. 의료 행위에 특화된 로봇의사는 보통의 인간의사 보다 월등한 지식과 능력 및 의료기술을 지닌 존재가 될 것이며, 건강과 장수를 희망하는 사람들의 요구에 편승해서 인간의사의 지위와 권위를 상당히 낮추게 될 것이다. 인간의사와 로봇의사의 격차는 더욱 벌어질 것이며, 그와 함께 그에 수반되는 의료윤리적 문제들은 더욱 많아질 것이다. 그와 함께 인간의사와 로봇의사 사이의 역할 관계도 역전될 가능성도 있다.

일반적 수준에서 로봇의사는 인간의사에 비교 우위를 가질 수 있지만, 반대로 심각한 결함을 지닐 공산은 크다. 목적 맞춤형 감정로봇의사의 가능성에도 불구하고 인간의사는 목적 맞춤형 로봇의사와는 비교할 수 없는 우위를 갖는데, 환자와의 의사소통과 대면진료에 있어서 환자의 내면적인 정신적 요인들에 대한 이해에 “심각한 결함”을 지닌 존재라 할 수 있다.¹⁴⁾ 환자에 대한 감정적 및 정신적 접촉과 이해는 임상의료적 측면에서 매우 중요하다. 로봇의사가 이런 부분에 대해서 결함을 갖고 있다는 것은 의사로서의 직무 수행에 엄격한 제한을 두어야 하며, 때문에 더더욱 인간의사의 지휘 감독과 통제가 필요할 것이다. 이러한 인식이 반영된 원칙 아래서 인간의사와 로봇의사의 협업이 이루어져야 한다. 따라서 우리는 인간의 생명과 건강 이라는 특정한 목적과 목표에 맞춰져 있는 로봇의사 역시 지식과 기술적 측면의 이성적 기능과 환자와 감정적 소통의 감성적 기능 모두에서 인간의사와 다를 바 없이 실수와 오류 가

14) P. Gelhaus, “Robot decisions: on the importance of virtuous judgment in clinical decision making”, 883-887쪽 참조.

능성을 전제로 가능한 윤리적 문제들을 고민해 보는 것이 옳을 것이다. 적어도 이런 차원에서의 의료 행위와 관련해서 로봇의사에게도 ‘마치 인간의사에게 요구하는 의료윤리의 원칙을 준수할 수 있는 것처럼’ 운영 및 관리되어야 한다. 역으로 이러한 윤리적 요구는 한편으로는 이러한 준(準) 윤리로봇의사이어야만 의사공동체의 일원으로서의 자격을 부여할 수 있으며, 다른 한편으로는 이렇게 인간의사의 지위에 준하는 지위를 가질 수 있는 로봇의사는 당연히 의료 행위의 의미와 목적에 적합한 의료 행위를 하는 책임 있는 존재여야만 한다.

의료윤리원칙은 기본적으로 의사의 소명과 책무 및 의사와 환자 사이의 관계를 기초로 하는 윤리적 원칙이다. 그러나 미래의 의료윤리는 양자의 관계만이 아닌 로봇의사와의 관계도 포함해야 한다. 그러나 로봇의사에게 히포크라테스 선서와 같은 전통적인 의사윤리지침의 정신과 원칙을 그대로 적용하기 어렵다. 로봇의사를 어떻게 볼 것인가, 즉 그 본성과 능력에 있어서 어떤 지위를 갖는 존재인가 하는 문제가 개입되어 있기 때문이다. 로봇의사의 존재 가능성은 우리가 어떤 인공지능을 장착한 로봇의사를 만들 수 있는지, 또는 만들어야 하는지 등의 근본적인 문제들을 던져 준다. 이러한 존재론적 문제들이 로봇윤리에 대한 고민을 수반하듯이 로봇의사에 대한 윤리적 이해를 요구한다. 인간의사의 지위와 행위에 대한 윤리적 요구처럼 로봇의사 또한 이와 유사한 책임과 의무를 가져야 하는 것은 분명하다. 이런 문제의식의 연장선상에서 강인공지능 내지는 초인공지능 수준의 로봇의사는 최소한 인간의사와 동등한 윤리적 행위자로 가정할 수 있다. 미래의 자율형 인공지능로봇 역시 목적과 문제를 스스로 설정할 수 있는 인간과 달리 특정한 목적에 맞게 조건지어진 존재라는 점에서 인간과 근본적인 차이를 갖는다 하더라도 의학과 의료라는 특정한 제한된 목적 아래 설계된 경우에는 적어도 의료 행위의 목적과 의도는 이해하고 있다고 보아야 하기 때문이다.

우리는 로봇의사가 활동하거나 운영하는 병원이 존재하는 세계에 살게

될 것이다. 실수는 인간의사만이 하는 게 아니므로 우리는 생존이나 치료 가능성 및 의무와 책임 등에 대해서 훨씬 복잡한 윤리적 문제 상황에 직면할 것이다. 특정한 목적 아래 설계된 인공지능 역시 인간처럼 실수와 오류를 완전히 피해갈 수는 없다. 그렇다면 인공지능 로봇의사의 오진이나 잘못으로 인한 책임은 누가 질 것인가? 아마 이런 문제가 발생할 수 있다는 것은 로봇의사 단독으로 이루어지는 의료 행위는 허용되지 않을 것이라는 것을 짐작케 한다. 이런 점들과 상황을 고려할 때 미래의 의료 행위는 인간의사나 로봇의사 단독으로가 아니라 인간의사와 로봇의사의 상호 협력 내지는 상호 보완적인 작업과 관계 하에 이루어질 것으로 보아야 한다.

지난 2018년 4월에 “당뇨성 망막병증을 ‘의사의 개입 없이’ 진단하는 IDx-DR이라는 인공지능을 허가”한 사례가 보도된 바 있다.¹⁵⁾ 이는 어디까지나 현재 상용되고 있는 혈압측정기나 한방 부황기기처럼 개인적 용도의 의료기기에 해당한다고 볼 수 있다. 이에 준하는 인공지능 메디컬/헬스케어 의료기기들의 개인적 사용은 엄밀한 의학적 기준에 기초해서 허용되도록 관리 통제가 이루어져야 한다.¹⁶⁾ 반대로 인간의사의 승인과 감독이 필요한 의료행위에 대해서는 철저한 관리가 이루어져야 하며, 로봇의사 역시 이러한 대상에 속한다. 미래의 의사윤리 및 의료윤리의 원칙은 로봇의사의 의료 활동을 인간의사의 책임과 감독 아래 이루어지도록 제한해야 할 것이며, 이러한 관점에서 인간의사와 로봇의사의 관계와 역할을 반영하는 것이 중요한 과제가 될 것이다. 따라서 이에 대해서도 새로운 인식 전환이 요구된다. 이에 현재 우리나라의 대한의사협회의 ‘의사윤리강령’과 ‘의사윤리지침’에 준하는 미래의 로봇의사에게 필요한 윤리 준칙을 마련할 필요가 있다. 때문에 나는 기존의 인간중심의 의사윤리가 아니라 로봇의사의 관점에서 인간의사와 로봇의사의 관계가 반영된

15) 최윤섭, 『의료 인공지능』, 395쪽.

16) 변순용·송선영, 『로봇윤리란 무엇인가?』, 130-131쪽 참조.

로봇의사윤리지침을 제안하려 한다.

4. 로봇의사의 윤리원칙

나는 인간의사와 비교할 때 기술형 및 목적형 로봇의사를 지식과 정보의 우위를 지닌 존재로서 규칙과 원칙에 의거한 의료 직무 수행에 특화된 의사로 상정한다. 그리고 인간의사 수준의 자율적 행위자는 아니기 때문에 상황윤리적 판단이나 배려윤리적 행위에는 제약이 따르는 비인간적 인격체로 간주한다. 이러한 조건과 기준에 의하면, 로봇의사는 “규칙과 원칙에 기반한” 윤리 규범을 준수하는 행위 지향을 갖는다고 할 수 있다.¹⁷⁾ “윤리로봇(ethicbots)”으로도 불릴 수 있는 로봇의사 혹은 윤리로봇의사(ethicbot doctor)는 ‘의료’라는 특수한 목적에 특화된 존재이기 때문에 의료 이외의 문제 상황에 대처하는 데에는 근본적인 한계를 가질 것이다.¹⁸⁾ 그리고 무엇보다도 인공지능이 인간과 같은 자연지능 수준에 이르기 위해서는 자기생산체계에 기반을 두는 존재이어야 하지만,¹⁹⁾ 나는 인공지능은 자기 목적 정립적 존재 혹은 자율적 문제 발견과 설정의 주체가 될 수는 없다고 생각한다. 다만 인공지능적 존재의 최대한은 특수한 문제 해결에 특화된 기능형 및 목적형 인공물이며, 그것의 한 가지 구체화가 로봇의사일 수 있다고 생각한다. 다시 말해서 인간의사의 자율적 판단과 행위규칙을 로봇의사에게 완벽하게 그대로 프로그래밍 하거나 구현할 수 있는 인공물의 창조는 불가능할 것이며, 이러한 제약과 한계가 있는 로봇의사 역시 진단과 치료법의 선택에서 오류와 실수를 저지를

17) P. Gelhaus, “Robot decisions: on the importance of virtuous judgment in clinical decision making”, 883쪽.

18) “윤리로봇”이라는 용어에 대해서는 다음을 참조. 라파엘 카푸로 & 미카엘 나겐 보르그 편저, 『로봇윤리』, 19쪽.

19) 박충식, 「생명으로서의 인공지능: 정보철학적 관점에서」, 34-39쪽.

수 있는 존재이며, 따라서 인간에 의한 관리와 통제가 필요한 존재이다.

이처럼 장차 구현될 강인공지능 혹은 초인공지능 수준의 로봇의사는 자율적인 목적과 문제의 설정이라는 인간 고유의 능력을 갖출 수는 없기에 인간의사와 동일시 할 수 없다. 그러나 반대로 이미 인간의 생명과 질병 및 건강이라는 특정한 목적과 그와 관련한 문제 설정이 프로그래밍되어 있는 로봇의사는 의학과 의료라는 특수한 제한된 분야에서는 인간 의사를 능가할 공산이 크다. 어쩌면 수많은 인간 의사들의 총합 이상일 수도 있는 로봇의사를 인간 의사와 비교하는 것 자체가 무리일 수도 있다. 그러나 윤리적 문제는 이런 기능적 우월성과는 상관없이 존재한다. 이러한 이유에서 나는 의사공동체의 일원이 될 로봇의사의 존재 지위와 의료행위의 자격 조건을 마련하는 것을 매우 시급한 과제로 인식한다. 이 일은 역으로 우리가 어떤 로봇의사를 만들 것인지에 대한 윤리적 요구를 제시하는 것이기도 하다. 한마디로 로봇의사의 윤리원칙의 필요성은 모든 로봇의 존재 조건과 자격을 제한하는 로봇윤리의 필요성을 실증하며, 그 구체화의 실례가 될 수 있다. 더욱이 인간의 생명과 건강이라는 가장 중요한 가치를 지닌 영역과 관계한다는 점에서 인공지능 로봇의사의 의료윤리 지침은 로봇윤리 전반에 중요한 기여를 하게 될 것이다.

나는 로봇 의료윤리의 필요성과 관련해서 일련의 <로봇의사윤리강령>과 <로봇의사윤리지침> 등 의료행위 일반에 공히 적용할 수 있는 가장 일반적 원칙으로서 최소한의 로봇의사가 지켜야 할 윤리 준칙을 제안하고자 한다. 이 준칙은 기본적으로 인간 의사 및 환자와의 사이에서 지켜야 할 윤리원칙이다. 나는 오래전에 미국의 생명윤리 학자들인 비침과 칠드리스가 제안한 바 있는 생명의료 윤리의 네 가지 원칙, 1) 자율성 존중의 원칙 2) 악행 금지의 원칙 3) 선행의 원칙 4) 정의의 원칙 등을 준용한 변형 원칙으로서 ‘로봇의사윤리원칙’을 제안할 것이다.²⁰⁾ 그러나

20) Tom L. Beauchamp and James F. Childress, *Principles of Biomedical Ethics*, 3/4/5/6절 참조.

이들이 제시한 생명의료윤리의 원칙들과 관련해서는 세부적으로 많은 수용과 제한 및 비판적 검토들이 이루어져왔으며, 이 원칙들의 과연 생명의료윤리적 문제들을 해결할 수 있는 기능을 제대로 할 수 있는지, 또 원칙들 간의 충돌을 어떻게 해결할 수 있는지 등 이 원칙들에 대한 논의들이 꾸준히 제기되어 왔다.²¹⁾ 그럼에도 나는 비침과 칠드리스의 4원칙주의 생명의료윤리를 미래의 로봇의사의 윤리원칙으로 변용 · 확대함으로써 내가 제안하는 윤리원칙들이 환자와의 관계에서 결코 분리될 수 없는 인간 의사와 로봇 의사의 지위와 역할에 대한 윤리적 논의들로서 연속성을 갖고 다루어지기를 기대한다. 비침과 칠드리스의 생명의료윤리의 원칙들은 “직견적으로 타당하지만 절대적으로 타당하지 않은”²²⁾ 준칙들이라 할 수 있지만, 이 또한 로봇의사윤리의 원칙들을 정립하는 가운데에서 구체화 및 체계화될 수 있을 것이다.

이 미래의 로봇의사윤리의 원칙은 로봇의사가 의사공동체의 일원으로서 받아들였을 때, 현재의 인간중심의 의료윤리의 원칙에 덧붙여 환자 및 인간 의사와의 관계를 중심으로 로봇의사의 관점에서 마련한 준칙들로 구성된다. 따라서 이 준칙은 일차적으로 로봇의사의 윤리원칙이지만, 그 내용은 환자 및 인간 의사와의 관계에 집중한 것이기에 인간 의사와 로봇 의사가 동시에 그리고 함께 준수해야 할 윤리 지침이기도 하다. 아울러 인공지능 로봇의사는 최소한 이러한 준칙에 따라 행위 할 수 있는 존재이어야 하며, 또 그렇게 프로그래밍 되어야 한다. 그 경우에만 의사로서의 자격을 가져야 한다. 향후 모든 의료 행위에 로봇의사의 참여가 필수적이며, 또한 로봇의사 단독으로 독자적인 의료행위를 허용할 수 없

21) 이와 관련해서는 다음을 참조. 김종국, 「4원칙주의 생명의료윤리」와 칸트 도덕철학, 『칸트연구』 제37집, 2016; 이상목, 「의사의 책임윤리의 토대로서 선행과 신뢰」, 『철학논총』 제77집, 새한철학회, 2014; 최경석, 「생명의료윤리에서의 ‘자율성’에 대한 비판적 고찰」, 『한국의료윤리학회지』 14(1), 한국의료윤리학회, 2011.

22) 김종국, 「4원칙주의 생명의료윤리」와 칸트 도덕철학, 21쪽.

는 한, 로봇의사는 인간의사의 협업과 통제 아래 의료 행위를 해야 한다. 그러므로 로봇의사의 윤리원칙은 동시에 보편화된 인공지능 시대의 모든 의사들의 윤리원칙이라 할 수 있다.

이 로봇의사의 윤리원칙에 비견할만한 것으로서는 자신의 소설 <술래잡기>(Runaround, 1942)에서 묘사한 바 있는 아시모프(Isaac Asimov)의 “로봇공학 3법칙”(Three Laws of Robotics)이 있다. 그 각각을 “인간보호”, “명령복종”, “자기보호”의 원칙으로 별명할 수 있는 3개의 법칙은 다음과 같다.²³⁾

법칙1: 로봇은 인간에게 해를 입히지 않아야 한다. 그리고 위험에 처한 인간을 외면해서도 안 된다.

법칙2: 제1원칙에 위배되지 않는 한, 로봇은 인간의 명령에 복종해야 한다.

법칙3: 제1원칙과 제2원칙에 위배되지 않는 한, 로봇은 자신을 지켜야 한다.

그리고 나중에 아시모프는 소설 <로봇과 제국>(1985)에서 ‘법칙0’라 할 수 있는 네 번째 법칙을 추가했다. 그가 다른 세 법칙들이 위배할 수 없는 대원칙으로서 제시한 법칙0은 다음과 같다.

법칙0: 로봇은 인류에게 해를 가하거나, 행동을 하지 않음으로써 인류에게 해가 가도록 해서는 안 된다.

이처럼 소설적 구상으로 제시한 아시모프의 3법칙은 기본적으로 인간을 위한 로봇을 만들라는 요구를 담고 있다. 이러한 정신은 로봇의사에게도

23) 이와 관련한 논의들에 대해서는 다음을 참조. 웬델 윌러치, 콜린 알렌, 『왜 로봇의 도덕인가』, 12-13쪽; 변순용·송선영, 『로봇윤리란 무엇인가?』, 58-64; 신현우, 「진화론적 지능형 서비스 로봇에 대한 실천윤리학적 고찰」, 4-5쪽.

적용되어야 한다. 그런 점에서 내가 제안하는 로봇의사의 윤리원칙 역시 아시모프의 ‘3법칙’을 의료윤리적 관점에서 재해석하고 구체화한 것이라 할 수 있다. 나는 아시모프의 로봇공학 3법칙이 윤리로봇의 필요성과 당위성의 기본정신을 아주 간단명료하게 표현하고 있다고 생각한다. 또한 이러한 통찰력 안에는 진심으로 심각한 윤리적 고민들도 함께 묻어 있다는 것을 느낀다. 그것은 우리가 구체적으로 고민해 보지 않으면 안 되는 윤리적 문제들이다. 과연 윤리로봇이 가능할 수 있는지, 한 인간 혹은 인류가 처한 복잡한 상황에 대한 객관적 판단을 할 수 있는지, 판단과 결정의 올바른 기준을 어떻게 선택할 수 있는지, 또 기준은 누가 마련해줄 것이며, 과연 그러한 기준에 대한 일치된 견해가 가능할 수 있는지 등등 현재 우리가 안고 있는 적지 않은 철학적 및 윤리적 문제들을 고스란히 짊어지고 있는 로봇의 모습이 그려진다. 내가 가정하듯이 비록 건강 증진과 질병 치료라는 행위 목적과 문제 상황에 제약되어 있다고 하더라도 로봇의사 역시 이와 유사한 문제들을 안고 있으며, 머지않아 그것은 현실이 될 공산이 크다.

1) 자율성과 상호 동의의 원칙

로봇의사윤리의 첫 번째 원칙은 ‘자율성과 상호 동의의 원칙’이다. 이 원칙은 현재의 인간 의사 중심의 ‘자율성 존중의 원칙’을 변용한 것이다. 이 자율성 존중 원칙의 근본정신은 ‘환자의 자율적 의사에 따라서 진료 행위를 해야한다’는 것이다. 다시 말해 환자의 자율성을 존중하는 원칙으로서 진단과 치료 선택에서 타인이나 외부의 간섭 없이 환자 스스로 의료행위를 결정하는 권리를 보호하는 원칙이다. 이는 기본적으로 치료를 위한 환자 자신의 자율적 의사를 중시한다. 즉 의사는 환자의 질병 및 치료의 필요성에 대한 충분한 설명을 통해 환자의 동의를 득한 후에 치

료를 개시해야 한다. 의사의 환자에 대한 의무로서 “충분한 설명에 의거한 동의(informed consent)”²⁴⁾의 준칙에 따라서 환자의 자율적 선택을 돕기 위해서 의사는 환자의 질병에 대한 최대한의 또는 충분한 정보를 제공하여야 하며, 그 도움은 어떤 편견이나 오해의 소지가 없는 객관적이며 중립적인 것이어야 한다. 말하자면 치료의 선택이나 거부를 유도하거나 의도를 지닌 정보 제공은 비윤리적인 행위이다.

의료적 의사결정에 로봇의사의 참여는 우리가 가장 중요하게 고려해야 하는 문제다. CPU의 처리 능력이 인간의 뇌 수준으로까지 상승할 2025년쯤에는 아마 대부분의 의사결정 과정에 인공지능의 참여가 당연시 될 공간이 크다. 실제로 이미 이와 같은 일이 벌어지고 있는데, “생명과학, 암 연구, 노화에 따른 질병, 재생의학 분야에 투자하는 홍콩의 투자 금융회사 딥날리지 벤처스는 바이탈(VITAL, Validating Investment Tool for Advancing Life Sciences)이라 불리는 인공지능 알고리즘을 이사로 임명했다.”²⁵⁾ 이는 의사결정자에게 중요한 정보를 제공하는 행위만이 아니라 주어진 문제 해결에 대한 최선의 결정을 내리는 역할을 포함한다. 이러한 역할과 직무 수행은 그것이 보조 혹은 협업이든 인간 의사와 로봇 의사 모두에게 요구된다. 환자에게 가감 없이 최고의 정보를 제공하고 환자가 최선의 결정을 하도록 돕는 것이 의사의 의무이기 때문이다.

로봇의사는 환자의 의사결정을 돕는 일을 자기 혼자 해서는 안 된다. 로봇의사의 진단 및 치료에 관한 정보 혹은 판단은 언제나 사전에 인간 의사의 동의를 거쳐야 한다. 오류와 실수 가능성을 고려할 때, 의료 행위에 특화된 로봇의사일지라도 무엇이 최선의 치료인지에 관한 판단은 인간 의사와의 의사결정 과정을 거쳐야 한다. 따라서 진단과 치료에 관한 환자의 동의를 얻기 위한 의사의 설명 의무는 인간 의사와 로봇 의사 모두

24) Tom L. Beauchamp, and James F. Childress, *Principles of Biomedical Ethics*, 62-82쪽.

25) 클라우스 슈밥, 『제4차 산업혁명』, 219쪽.

에게 있지만, 최종 결정의 책임은 인간의사의 몫이다.

인간의사는 로봇의사를 의사결정의 상대자 내지는 협력자로 대해야 하며, 이를 전제로 로봇의사는 의사결정에 참여해야 한다. 더욱이 개인이 쉽게 조작할 수 있는 간단하고 위험이나 부작용이 거의 없는 첨단 가정용 의료진단 기기들은 이미 상용되고 있듯이 언젠가는 로봇의사 단독으로 진단과 치료가 허용되고, 그 영역과 범위는 더 확대될 수 있다.²⁶⁾ 이 경우에 로봇의사는 윤리로봇으로서 ‘로봇의사의 윤리원칙’의 준수 가능성이 자격조건이 되어야 할 것이다. 그리고 윤리원칙들을 보다 하위 준칙으로서 세부적 실행 규칙이 만들어져야 할 것이다. 이 경우에 가장 중요한 부분이 바로 의사소통 및 의사결정 과정을 준수할 매뉴얼을 마련하는 것이다. 윤리적 의사결정을 위한 매뉴얼을 프로그래밍 할 수 있느냐 하는 문제가 있지만, 이 또한 다른 분야와 마찬가지로 생명과 건강이라는 특수 목적에 적합한 인공지능의 존재에 비례해서 그 준칙 매뉴얼을 정교화하면 할수록 그 실현 가능성도 높아질 것이다.²⁷⁾ 다른 한편으로 이 같은 도덕적 조건 아래서 이루어지는 인간의사와 로봇의사의 공동의사결정은 넓은 의미에서 “도덕적 의사결정 기계”를 운용하는 것과 같은 일이 될 수 있다.²⁸⁾

환자에 대한 설명 의무와 동의 의무의 원칙 준수에 있어서 로봇의사와 환자 사이에는 간과할 수 없는 의사소통의 윤리가 적용된다. 환자의 동의는 누가 구할 것인지, 즉 인간의사와 로봇의사 중에 누가 환자의 동의

26) R&D정보센터, 『헬스케어 4.0시대 국내외 의료산업 패러다임과 신개념 의료기기 트렌드』, 지식산업정보원, 2018. 헬스케어산업 & 의료산업 참조.

27) Jatinder N.D. Gupta, Guisseppi A. Forgionne, Manuel Mora T (eds.), *Intelligent Decision-making Support Systems: Foundations, Applications and Challenges*, Springer, 2006 참조. 인공지능 ‘의사결정 지원체계’(DMSS)는 개인이나 조직의 의사결정과정을 지원하는 컴퓨터기반체계다. 최근에 정보공학과 인공지능 분야에서 이루어지고 있는 성과들은 인공지능-DMSS를 향상 및 발전시키고 있다.

28) 이에 대해서는 웬델 윌러치, 콜린 알렌, 『왜 로봇의 도덕인가』, 19, 67-95쪽 참조.

를 구하는 관련 정보를 제공하며, 적절한 치료 방법과 치유 가능성 등에 대해서 설명할 자격을 가지며, 어떻게 소통할 것인가? 가령 인간의사와 로봇의사의 진단에 차이가 있을 수 있으며, 이 경우 어떤 소견을 선택해야 하는 문제가 생길 수 있으며, 그 최종 결정은 인간의사의 몫이어야 한다. 하지만 여전히 문제는 남는다. 환자에게 로봇의사와 인간의사의 치료법이 상이하거나 다른 치료법에 따른 결과가 엇갈리는 상황이 발생할 때, 우리는 그 최종 결정을 환자의 자율적 선택에 맡겨야 하는가? 아니면 전문적인 지식이 없는 환자를 대신해서 인간의사 혹은 로봇의사가 결정할 것인가? 인간의사의 결정에 따른다 하더라도 그 선택지가 로봇의사의 진단일 경우, 과연 환자의 동의를 얻기 위한 책임자는 누구여야 하는가? 어떤 점에서는 로봇의사 역시 진단과 치료의 주체라고 한다면 의당 환자의 동의를 얻는 행위는 로봇의사의 몫이어야 할 것이다. 그러나 과연 그럴 수 있으며, 또 왜 그래야 하는가 하는 것은 여전히 문제다. 그리고 만일 선택의 결과에 따른 책임은 누가 져야 하는가? 이런 문제들을 해결하기 위해서 인간의사와 로봇의사의 공동 의료 행위에 적용할 수 있는 윤리원칙이 필요하다. 나는 이렇게 로봇의사를 포함하는 의사공동체의 의료행위에 있어서 로봇의사의 진단을 선택하거나 치료를 포함한 인간의사의 대행역할에 따른 책임은 누가 져야 하는지 등 일련의 예상되는 문제들과 관련해서 결국 환자, 인간의사, 그리고 로봇의사의 상호 동의가 진단과 치료 결정의 원칙이 되어야 한다고 말할 것이다.

그런데 인간은 자신의 삶을 스스로 결정할 권리가 있다.²⁹⁾ 이러한 자신의 결정에 반하는 행위를 강제할 상위의 권리가 있을 수 없기 때문에 인간의사와 로봇의사의 상호 동의에 의거한 진단과 치료의 결정 보다 환자의 자율성 원칙이 더 우선해야 한다. 환자의 최종 선택이 이루어지는 과정에서 인간의사는 의사 상호 간의 의사결정 과정에서 논의되었던 진

29) 이와 관련해서는 맹주만, 「칸트와 미학적 자살」, 2015; 「안락사와 존엄사, 그리고 웰다잉법」, 2016 참조.

단과 치료에 관한 진실을 환자에게 알려주어야 하며, 이에 근거해서 환자의 최종 결정이 이루어지도록 해야 한다. 나는 이렇게 환자의 자율성을 존중하고 그와 관계있는 인간 의사와 로봇 의사의 사전 동의에 의거한 의료 행위를 ‘자율성과 상호 동의의 원칙’이라 부른다. 자율성이 무엇인지 완벽하게 정의하는 문제는 어려운 철학적 작업이 될 것이다. 나는 다만 ‘자율성 존중의 원칙’에서 존중(respect)이라는 표현을 제거한, 보다 강한 의미의 자율성 개념을 사용한다. 존중이라는 표현은 중시한다, 고려한다 등의 의미를 담고 있는데, 이에는 무조건적인 당위성 개념이 희소하다. 나는 단적으로 자율성을 가장 중요한 원칙으로 채택하려 한다. 그것은 반드시 지켜져야 하는 윤리원칙이다. 뇌사자나 미성년자의 경우처럼 이를 준수할 수 없는 예외적인 경우에는 별도의 준칙을 마련하면 될 것이다. 나는 로봇 의사가 활동하게 될 미래의 의료윤리에서는 이 점이 더욱 더 강조되어야 한다고 믿는다.

이 첫 번째 윤리원칙은 정상적인 의사표명과 판단이 가능한 환자는 물론이고 뇌사상태에 있어서 의사 표명이 불가능한 환자, 정신질환자여서 정확한 판단과 의사표명에 문제가 있는 환자, 두려움과 공포, 혹은 자신의 가치관으로 인해서 치료를 거부하거나 그럴 가능성이 높은 것으로 인지된 환자 등의 치료와 관련한 모든 경우에 자율성 원칙의 일시적인 유보 혹은 거짓말의 허용 가능성을 놓고 고민하는 문제에도 적용할 수 있다. 가령 정상적인 환자의 경우에 통상 거짓말의 유용성을 놓고 고민하는 것은 진실을 말하는 것이 환자의 자율적 결정이나 치료에 대한 부정적 영향을 우려하기 때문이다. 그런데 이러한 경우에 그 원인의 일부는 의사가 하는 ‘충분한 설명’의 불충분함이나 불확실성에 있을 수 있다. 즉, 환자에게 동의를 구하기 위해서 하는 ‘충분한 설명’이 실질적인 정보력을 갖지 못하는 때와도 관련이 있다. 의사에 대한 신뢰도가 높을수록 환자의 의사결정은 더 바람직한 결정을 내릴 수 있는데, 인간 의사 대비 로봇 의사에 대한 신뢰도는 더욱 높아질 것으로 예상되기 때문이다.

환자가 의사의 진단과 치료 가능성을 듣게 될 때, 의사가 확신을 갖고 환자에게 충분한 설명을 하는 것도 쉽지 않다. 이럴 경우 듣게 되는 진실은 환자들의 심정에도 마찬가지로 불안과 망설임이 생길 수 있다. 향후에는 환자 혹은 대리인은 의사공동체의 일원이 된 로봇의사의 진단의 정확성과 치료 가능성에 대한 설명을 더 높이 평가하거나 신뢰할 확률이 높다. 더욱이 인간의사와 공동으로 진행되는 의학적 판단의 경우는 그 신뢰도는 더욱 높아질 것이다. 이렇게 되면 환자의 자율성과 거짓말의 불필요성의 상관성은 훨씬 높아질 것이라 예상할 수 있다.

2) 악행 금지와 공동 책임의 원칙

인간의사 중심의 악행 금지의 원칙은 ‘환자에게 해악을 입히거나 환자의 상태를 악화시키는 의술을 사용해서는 안 된다’는 것이다. 히포크라테스 선서에 적시되어 있듯이 의술은 선한 일에 쓰여야 하며, 반대로 해를 입히는 일에 사용되어서는 안 된다. 전자가 선행의 원칙을 대변한다면, 후자에 해당하는 것이 악행 금지의 원칙이다.

자율성과 상호 동의의 원칙에서 인간의사와 로봇의사의 의견 불일치 경우와 마찬가지로 환자 자신의 이익과 해악을 고려해서 결정되어야 할 의료 행위에서 인간의사와 로봇의사의 결정이 상이하거나 차이가 있을 수 있다. 이는 의료법의 차원을 넘어서 인간의사들 사이에서도 선택이 갈리기도 하는 문제이기도 하다. 수술 치료가 어떤 결과를 가져올지 장담할 수 없는 경우나 말기암의 경우처럼 환자를 위해서 치료를 계속해야 하는지, 아니면 극심한 고통을 야기하면서도 불필요한 연명치료에 지나지 않는지 등은 보기에 따라서는 악행 금지의 원칙의 준수 혹은 위반 행위로 갈릴 수가 있다.

그러나 악행 금지에서의 ‘악행’이 무엇인지를 정의하기는 쉽지 않은

철학적 문제를 수반한다. 만일 ‘악행 금지’(nonmaleficence)가 단순히 고통을 야기하거나 증가시키는 행위를 대상으로 하는 것이라면 의사가 환자를 위해 할 수 있는 조치는 상당히 제약된다. 질병을 치료하는데 있어서 고통을 수반하지 않는 경우는 거의 없을 것이기 때문이다. 그러므로 악행을 어떻게 정의하느냐에 따라서 인간 의사와 로봇 의사의 판단은 극명하게 대립될 수 있다. 인간 의사와 달리 로봇 의사는 어떻게 프로그램 되느냐에 따라서 고통의 의미를 이해하는데 상당한 제약이 따르거나 잘못 이해할 수 있다. 이는 로봇 의사의 존재론적 지위와 관련하여 심각한 문제를 야기할 수 있으며, 때문에 로봇 의사의 역할을 상당히 제한해야 할 수도 있다.

의료적 문제에 있어서 악행 금지의 원칙은 대체로 ‘~을 하지 말라’는 형식으로 제안되는 특징이 있다. 물론 이는 윤리적 의무와 법적 의무의 기초에 인간의 자율적 의사가 더 중요한 요소로 작용하기 때문에 생기는 언어적 문제일 수도 있다. 가령 일반적으로 ‘선행을 베풀어라’고 말하며, 또 ‘나쁜 짓은 하지 마라’고 말하지 반대로 표현하지는 않기 때문이다. 하지만 ‘더 이상 치료를 하지 마라’와 같은 표현의 경우처럼 치료를 하는 것이 곧 나쁜 행위가 되는 것은 아니듯이 이 같은 어법을 일반화시킬 수는 없다.

악행의 대상, 즉 피해를 입는 환자의 입장에서 보면, 로봇 의사는 질병의 치료를 기능적으로 이해한다고 볼 수 있는데, 이 경우의 로봇 의사는 저와 같은 치료의 선택 역시 기능적으로 할 것이며, 이는 전적으로 질병에 대한 정의, 즉 인간 신체의 건강상태와 관련해서 프로그램 된 정상과 병변의 구분에 근거해서 진단을 할 것이다. 그런데 치료 가능성의 기능적 선택은 전적으로 환자만을 대상으로 내리는 결정이므로 이것만으로도 그 시행 여부를 결정할 수는 없다. 가령 장기 이식 대상자가 있을 때와 없을 때에 따라서 수술 여부의 중단 등의 결정이 달라질 수 있으며, 따라서 다른 방식의 치료를 결정해야 할 수도 있을 것이다. 이런 문제 때

문에 실질적으로 치료 중단 혹은 대체 여부의 최종 결정은 고통의 의미를 이해하고 있는 인간 의사의 몫이 될 것이다. 그러나 장기적 치료의 전망에서 병변의 상태를 보다 정확히 진단할 수 있으면, 그에 수반되는 고통의 정도가 보다 정확히 예측할 수 있을 것이기 때문에 로봇 의사가 담당하는 신체 구조와 조직에 대한 보다 엄밀한 생리학적 이해와 그 각각의 증상에 대한 물리적 반응의 정확한 측정 가능성은 치료 결정에 도움이 될 것이며, 또한 고통의 정도에 대한 계산 가능성도 더 높아진다면, 도움이 될 것이다. 더욱이 현재의 인간 의사의 진단과 처방시에 내리는 치료 가능성 및 고통 등에 대한 예상과 조언은 의사마다 약간씩 다르기도 하지만 상당히 대략적인 경향이 있다. 이런 문제에 있어서 로봇 의사의 보조는 매우 의미 있는 역할을 하게 될 것이다. 즉, 이러한 결정에 로봇 의사는 중요한 역할을 담당할 것이므로 악행 금지의 원칙의 정신에 비추어볼 때 특히 그 결정에 있어서 인간 의사와 로봇 의사는 공동의 책임을 져야 하며, 이러한 방식으로 양자의 협업의 이루어져야 할 것이다. 이에 나는 로봇 의사의 두 번째 윤리원칙을 ‘악행 금지와 공동 책임의 원칙’으로 부른다.

의사공동체의 일원으로서의 지위와 역할에 맞게 책임을 다하려면 악행 금지의 원칙에 준하는 의사로서의 책임과 의무는 물론 인간 의사의 결정에 도움이 되어야 할 것이다. 그러므로 로봇 의사에게 공동 책임을 요구하는 것은 곧 그와 같은 수준의 로봇 의사를 제작할 것을 요구하는 것이기도 하다. 이에 기초해서 인간 의사와 환자는 치료의 선택에서 있어서 최선의 결정을 할 수 있게 될 것이다. 이를테면 인공 호흡기의 제거와 유지 가운데 어느 것이 환자에게 악을 행하는 결과를 낳는지에 대한 판단의 경우 금지된 진료와 해야만 하는 진료의 결정에 인간 의사와 로봇 의사는 공동의 책임을 지는 결정을 내려야 한다.

3) 선행과 돌봄의 원칙

인간의사 중심의 선행 원칙은 ‘의사는 타인의 질병을 치료하고 건강을 증진하도록 노력해야 한다’는 것이다. 악행과 마찬가지로 선행(beneficence)이 무엇인지 정의하기는 쉽지 않다. 돌봄 또는 배려 또한 마찬가지다. 선행을 하는 것은 곧 누군가를 배려하고 돌보아주는 일이지만 반드시 일치하는 것은 아니다. 눈에 보이는 않는 선행이나 배려도 있으며, 누군가에게 하는 선행이 반드시 그 사람에게 도움이 되는 것이 아닐 수도 있으며, 돌봄 역시 당사자의 의사에 반하는 불편한 행위가 될 수도 있다. 다만 그 의도에 있어서는 선행과 돌봄은 궁극적으로는 누군가의 행복을 증진시키는 행위라 할 수 있다.

일반적인 윤리적 관점에서 보면, “선행의 원칙은 일반적으로 타인을 돕는 적극적인 조치를 요구하기 때문에 악행 금지의 원칙보다 더 이타적이며 중요한 원칙이다.”³⁰⁾ 해악의 제거에 한정될 수 있는 악행 금지의 원칙에 비해서 “해악의 예방과 해로운 조건들의 제거 및 적극적인 선행”³¹⁾을 요구한다는 점에서 선행의 원칙은 삶의 질의 문제와 관계가 깊다. 무엇보다도 의료적 관점에서 타인의 선을 적극적으로 증진시킬 것을 요구하는 선행의 원칙은 특히 건강관리라는 측면에서 중요하다. 마치 부모가 자식의 건강을 위해서 예방의학에 각별히 관심을 갖는 것처럼 타인의 건강을 위한 적극적인 관여가 허용될 수 있는 윤리원칙이다. 이런 사항들을 종합적으로 고려하면 의료적 관점에서의 선행은 일종의 인간적 돌봄 혹은 배려의 원칙이라 할 수 있다. 인간의 일을 대신해서 더 잘 할 수 있는 기계를 만들려는 것이 인공지능로봇의 존재 이유라면, 의료윤리적 관점에서 로봇의사는 이 행복의 추구에 유익한 기여를 할 수 있는 존

30) Tom L. Beauchamp, and James F. Childress, *Principles of Biomedical Ethics*, 135쪽.

31) 같은 글, 같은 곳.

재가 되어야 한다.

그러나 인간으로서 환자가 생각하는 행복은 의사가 생각하는 행복과 다를 수 있듯이 선행 혹은 배려는 환자의 입장에 따라서 다를 수 있다. 의료적 건강함이 곧 삶의 행복과 동일시되는 것은 아닐 수 있기 때문이다. 선행의 핵심은 환자를 돕는 데 있다. 동일한 의미에서 돌봄의 윤리도 성립한다. 인간적 삶의 의미와 목적 그리고 인간의 마음에 대한 이해 등은 인간 의사도 접근하기 어려운 영역이다. 마찬가지로 의학적 견지에서도 건강한 삶을 정의하기는 어렵다. 생명과 건강에 특화된 혹은 전문화된 로봇의사라 할지라도, 마찬가지로 감정적 소통기능을 갖춘 감정로봇의사라 할지라도 특정한 감정에 맞춰진 소통을 문제 없이 해내기는 어려울 것이다. 그러므로 로봇의사는 이 ‘선행과 돌봄의 원칙’에 있어서 마치 인간의사처럼 해서는 안 된다. 로봇의사의 환자를 위한 선행과 돌봄은 전적으로 인간의사에게 조언하는 수준에서 제한되어야 한다. 이 점에서 로봇의사는 인간 혹은 인류에게 도움을 주는 존재로 만들어져야 하며, 특히 의료적 목적에 특화된 맞춤형 로봇이 아닌 인간적인 로봇의사로 제작되어야 한다. 의료적 목적에만 제한된 로봇은 인간적 삶의 깊이와 내면에 접근하기 어려울 것이며, 그만큼 환자를 위한 소통에는 큰 제약이 따를 것이기 때문이다. 마찬가지로 순 기능적인 의학적 판단에 그 역할을 한정해야 하며, 동시에 인간의사와의 상호 동의, 공동 책임의 윤리원칙에 더 충실해야 할 것이다.

4) 의료권과 정의의 원칙

인간의사 중심의 정의의 원칙은 ‘환자는 공정한 절차를 통한 치료를 받을 최소한의 의료권이 있다’는 것이다. 인간의 평등을 말하듯이 치료의 평등 역시 동등한 가치를 갖는다. 의료권은 “인간의 권리이지 특권이 아

니다.”³²⁾ 그렇다면 신분이나 지위 혹은 경제적 능력의 차이로 인한 치료의 불평등이 허용되어서는 안 된다. 그러나 이러한 윤리원칙은 현실적으로는 이상적인 원칙이다. 각 나라마다 의료법이 있듯이 의료 행위는 법과 제도의 지배를 받는 사회적 행위이며, 특히 얼마만큼의 질 좋은 치료를 받느냐 하는 것은 경제적 능력의 차이가 큰 영향을 미치는 것이 현실이다. 그러나 가능한 한 환자는 그 누구든 차별 없이 진료를 받을 권리가 있으며, 이를 정의의 원칙이라 부를 수 있다. 이런 점에서 기초해석하는 로봇의사윤리의 원칙들 중에서 네 번째 마지막 원칙을 ‘의료권과 정의의 원칙’이라 부를 것이다.

철학과 윤리의 관점에서 정의의 문제에는 다양하고 복잡한 쟁점들이 있듯이 의료 정의 역시 그러하다.³³⁾ 환자의 치료와 직접적인 관련이 있는 문제들만 살펴보아도 정부나 병원에서 제공되는 예산 지원을 위한 의료 연구 대상의 우선순위를 정하는 문제로부터 인공지능의 도입으로 발생하는 잉여의료자원분배 분배와 관련한 정당한 우선 치료권,³⁴⁾ 환자의 기대수명 대비 의료자원 분배 순위, 사회적 기여나 의료자원 개발비 후원자의 우선 치료권, 가족 대 가족의 장기 이식 우선권, 저소득층에 대한 의료서비스의 무상 지원, 최소한의 보건의료 및 그 할당과 분배의 도덕적 정당성, 치료의 평등권 수혜 대상의 범위 등에 이르기까지 많은 문제들이 존재한다. 그리고 이 모든 문제들과 관련해서 로봇의사의 존재는 의료 정의의 구현이라는 차원에서 중요한 기여를 할 수 있다. 특히 국민 건강 측면에서 예방의학의 구현을 그 예로 들 수 있다. 또 사회경제적 측면에서 로봇의사의 의료활동의 확산은 의사수의 부족이나 비용 문제와

32) Victor Sidel, “The Right to Health Care: An International Perspective”, 341쪽.

33) Tom L. Beauchamp, and James F. Childress, *Principles of Biomedical Ethics*, 168-198쪽.

34) H. Tristram Engelhardt, Jr. *Foundations of Bioethics*, 1996; 장운혁, 정창록, 「인공지능과 의료자원분배」, 397-403쪽.

밀접한 연관이 있는 시간적 공간적 제약을 개선할 수 있으며, 병원이나 인간 의사에의 접근이 용이하지 않거나 상당한 제약을 받고 있는 취약 계층에까지 의료 혜택을 제공할 수 있다. 보건과 공중의료 등에까지 로봇 의사를 도입하는 것은 적지 않은 비용이 들 것이다. 그러나 환자수 대비 로봇의사가 감당할 수 있는 범위는 상당히 광범위할 것이기 때문에 비용-이익 또는 비용-효율성 측면에서 충분히 감당할 수 있을 것이며, 투자 대비 수익의 임계점에 이르게 되면 로봇의사의 제작 비용은 더 저렴해질 것이며, 따라서 시간이 지날수록 더 많은 이익을 남길 것이다.

그런데 이처럼 장기적 관점에서 로봇의사의 보급으로 의료혜택의 수월성은 높아지겠지만, 이와 관련한 치료 비용은 증가될 수도 있다. 다시 말해서 진료와 진단의 신속함과 혜택과는 반비례하게 치료비용이 많이 소요되는 치료의 경우, 환자의 경제적 여건과 국가보험의 지원 등의 지원이 병의 시급성에 반하는 경우가 빈번하게 발생할 수 있기 때문이다. 이 또한 의료권 및 의료자원의 분배 문제에서 중요한 관건이 될 수 있다. 정의로운 의료제도가 제대로 운영 되지 않을 경우 부익부 빈익빈의 의료 부정의의 문제가 로봇의사의 일반적 보급이 이루어지는 시대에는 더 큰 악영향을 끼칠 수 있다. 더욱이 환자에게 장기 이식을 시행할 경우에도 부당한 차별이나 불공정한 절차로 인한 피해를 보는 일이 생길 수 있다. 가령 인공 심장이나 기타 장기의 이식이 필요한 환자들 중에서 수혜의 우선권은 누구에게 있으며, 어떻게 결정해야 하는가? 또 누가 누구에게 이식하느냐 하는 문제가 절차적 정당성을 확보하지 못한다면 더 큰 부정의가 발생할 것이다. 이에 의료윤리적 관점에서 모든 환자에게는 의료권이 있으며, 따라서 진단과 검진과 마찬가지로 누구든 치료에 있어서 부당한 대우를 받아서는 안 된다. 그러므로 이와 관련해서 로봇의사와 치료 자원의 정의로운 분배가 정당한 절차에 의해서 이루어질 수 있는 원칙과 기준이 마련되어야 한다. 장기이식과 이식 수혜자 우선순위를 결정하는 문제 등 다양한 의료적 문제들에 대해서 편견과 차별 없는 의료 행

위를 추구하는데 있어서 로봇의사는 가장 합리적인 선택이 가능한 위치에 있을 것이며, 그에 따라 이를 정의롭게 수행하는 것이 곧 네 번째 윤리원칙을 준수하는 일이 될 것이다.

5. 나오는 글

인공지능 로봇의사는 어떤 의사이어야 하는가? 과연 모든 면에서 인간 의사에게 요구하는 윤리원칙을 준수할 수 있는 로봇의사의 존재는 가능한가? 인간의 감정을 읽을 수 있는 인공지능은 가능한가? 행위 주체로서나 언어 주체로서나 이성과 감성을 함께 지닌 로봇의사는 가능한가? 나는 이글에서 이성과 감성을 모두 갖춘 인격적 존재로서 완벽한 자율형 로봇의사는 불가능하지만, 인간의 생명과 건강 증진이라는 특수한 목적에 따라 맞춤 제작된 로봇의사는 언어 주체 및 판단 주체가 될 수 있으며, 그런 점에서 의사공동체의 일원으로 간주되어야 한다는 전제 하에 미래의 로봇의사윤리의 원칙을 제안했다.

다양한 방식으로 제시되고 있는 현행 의료윤리의 원칙들은 인간중심의 윤리원칙으로서 기본적으로 의사와 환자 사이의 윤리적 원칙이다. 그리고 로봇의사가 등장하더라도 많은 세세한 부분에서는 현재의 의료윤리적 문제에 대한 해결 방식의 범주를 크게 벗어나지 못할 것이다. 그러나 중요한 부분에서는 단순한 협업 이상으로 언젠가는 피할 수 없는 인간 의사와 로봇의사가 함께 하는 의료 행위에는 단순한 취직전쟁과 같은 현실적 문제를 떠나서 해당 업종의 직무 능력과 관련한 갈등과 분쟁이 생길 가능성이 매우 높으며, 그 성패는 환자의 결정에 의해서 좌우될 것이다. 때문에 갈수록 로봇의사의 지위와 역할을 둘러싼 문제들을 심각하게 고민하게 될 것이며, 그에 따라 로봇의사의 지위와 역할을 둘러싸고 많은 윤리적 문제들이 제기될 수 있다. 그러므로 현재의 인간 의사 중심의 의료윤

리는 재고되어야 하며, 인간과 인공지능의 협업으로 이루어질 미래의 의료윤리는 인간 의사와 로봇 의사의 관계를 포함한 준칙들로 수정되어야 한다.

인간은 진단과 치료 과정에서 이해, 참여, 선택, 의지, 희망 등과 같은 인간성에 대한 존중이 포함되기를 원한다. 인간은 실존적 존재이기 때문에 병과 죽음, 진단과 치료에 있어서 희망, 절망, 미래에 대한 기대, 가족들에 대한 죄책감 등 각자가 지닌 삶의 비중과 무게에 따라서 상이한 반응을 보인다. 그러나 인간의 의사도 이런 인간의 내면세계를 감당하기가 어려운데, 그 이상의 감성마저 갖춘 로봇 의사의 출현을 기대하기는 더욱 어렵다.³⁵⁾ 인간의 의사와 로봇 의사 사이에는 넘을 수 있는 엄연한 간극이 존재한다. 인간의 마음은 복잡한 세계를 이루고 있으며, 최소한 환자에게 최적의 치료가 될 수 있는 것은 공감적 마음을 지닌 의사를 만나는 것이 될 것이다. 특수한 부문에 특화된 감정 로봇으로서의 로봇 의사도 일정 부분 도움이 될 수 있지만, 그런 점에서 더욱 더 인간의 의사와 로봇 의사의 역할 분담 및 전문화가 필요할 수 있다. 같은 이유에서 로봇 의사가 흉내 낼 수 없는 인간의 의사 중심의 의학 교육도 시급한 과제가 될 수 있다. 당연히 가장 인간적인 능력과 인격을 갖춘 유능한 의사가 최고의 의사가 될 것이다. 철학과 윤리로 무장한 인간 의사의 모습이 미래 의사의 가장 중요한 척도가 되리라 예상할 수 있다.

이 글에서 제안한 로봇 의사의 윤리 원칙은 향후 인공지능 로봇 의사의 등장과 함께 의료공동체에서 고민해야 하거나 고민하게 될 윤리적 문제들에 대한 이해를 돕기 위한 그 대강을 제시해본 것이다. 이 원칙들과 함께 그와 관련한 다양한 세부적인 문제들은 이제 본격적으로 검토되어야 한다. 이는 동시에 인간의 의사와 로봇 의사가 공존하게 될 미래의 의료공동체 내지는 의사공동체의 일원이 되기 위한 로봇 의사의 자격과 조건

35) P. Gelhaus, "Robot decisions: on the importance of virtuous judgment in clinical decision making", 886쪽 참조.

을 제시하고 또 제한하는 작업도 포함한다. 이러한 시도는 궁극적으로 우리가 어떤 윤리적인 로봇의사 즉 윤리로봇의사(ethicbot doctor)를 배출할 것인지에 대한 시급한 고민과 함께 선제적 대응이 요구된다는 것을 의미한다.

참고문헌

- 김종국, 「4원칙주의 생명의료윤리'와 칸트 도덕철학」, 『칸트연구』 제37집, 2016.
- 고인석, 「인공지능의 존재 지위에 대한 두 물음」, 『철학』 제136집, 2018.
- 곽재식, 『로봇 공화국에서 살아남는 법』, 구픽, 2016.
- 김종국, 「4원칙주의 생명의료 윤리와 칸트 도덕철학」, 『칸트연구』 제37집, 2016.
- 김재인, 『인공지능의 시대, 인간을 다시 묻는다』, 동아시아, 2017.
- 라파엘 카푸로 & 미카엘 나젠보르그 편저, 『로봇윤리』, 변순용 · 송선영 역, 어문학사, 2013.
- 레이 커즈와일, 『특이점이 온다』, 김명남 장시형 옮김, 김영사, 2017.
- 마틴 포트, 『로봇의 부상』, 세종서적, 2016.
- 맹주만, 「인간 복제와 인간의 가치」, 『철학탐구』 제12집, 2000.
- _____, 「인간의 본성과 유전자 조작」, 『철학탐구』 제14집, 2002.
- _____, 「칸트와 미학적 자살」, 『칸트연구』 제36집, 2015.
- _____, 「안락사와 존엄사, 그리고 웰다잉법」, 『철학탐구』 제44집, 2016.
- 박영숙 제롬 글렌, 『세계미래보고서 2055』, 비즈니스북스, 2017.
- 박충식, 「생명으로서의 인공지능: 정보철학적 관점에서」, 『인공지능의 존재론』, 한울아카데미, 2018.
- 변순용 송선영, 『로봇윤리란 무엇인가?』, 어문학사, 2015.
- 신상규, 「인공지능은 자율적 도덕행위자일 수 있는가?」, 『철학』 제132집, 2017.
- 신현우, 「진화론적 지능형 서비스 로봇에 대한 실천윤리학적 고찰」, 『윤리연구』 Vol. 79, 2010.
- 웬델 윌러치, 콜린 알렌, 『왜 로봇의 도덕인가』, 메디치, 2014.
- 이상목, 「의사의 책임윤리의 토대로서 선행과 신뢰」, 『철학논총』 제77집, 새

- 한철학회, 2014.
- 장운혁, 정창록, 「인공지능과 의료자원분배」, 『한국의료윤리학회지』 제20권 제4호(통권 제53호), 한국의료윤리학회, 2017.
- 제러미 하워, 『증거기반의학의 철학 : 의학의 새로운 패러다임』, 전현우 역, 생각의힘, 2018.
- 천현득, 「인공지능은 감정을 가질 수 있을까?」, 『인공지능의 존재론』, 한올아카데미, 2018.
- 최경석, 「생명의료윤리에서의 ‘자율성’에 대한 비판적 고찰」, 『한국의료윤리학회지』 14(1), 한국의료윤리학회, 2011.
- 최윤섭, 『의료 인공지능』, 클라우드나인, 2018.
- 클라우드 슈밥, 『제4차 산업혁명』, 메가스터디(주), 2016.
- R&D정보센터, 『헬스케어 4.0시대 국내외 의료산업 패러다임과 신개념 의료기기 트렌드』, 지식산업정보원, 2018.
- Asimov, I., *I, Robot*, New York: Spectra, 2008.
- Beauchamp, Tom L. and Childress, James F., *Principles of Biomedical Ethics*, Oxford University Press, 1979.
- Boden, Margaret A. (ed.), *The Philosophy of Artificial Intelligence*, Oxford University Press, 1990.
- Bostrom, N., *The Transhumanist FAQ*, www.nickbostrom.com (Version 2.1 2003)
- _____, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.
- _____, “In defense of Posthuman Dignity”, *Bioethics*, Vol. 19, No. 3, pp. 202-214, www.nickbostrom.com, 2005.
- Edelman, Gerald M., *Second Nature: Brain Science and Human Knowledge*, Yale University Press, 2006.
- Engelhardt, Jr. H. Tristram, *Foundations of Bioethics*, Oxford University,

- 1996.
- Gelhaus, P., “Robot decisions: on the importance of virtuous judgment in clinical decision making”, *Journal of Education in Clinical Practice* 17 (2011) 883-887, Blackwell Publishing Ltd.
- Jatinder N.D. Gupta, Guisseppi A. Forgionne, Manuel Mora T (eds.), *Intelligent Decision-making Support Systems: Foundations, Applications and Challenges*, Springer, 2006.
- Kass, L. R., “Ageless Bodies, Happy Souls: Biotechnology and the Pursuit of Perfection”, www.TheNewAtlantis.com, 2003.
- Oliver, Richard W., *The Coming Biotech Age*, New York: McGraw-Hill, 2000.
- Rifkin, J., *The Biotech Century*, New York: Jeremy P. Tarcher/Putnam, 1998.
- Russell, S. and Norvig, P., *Artificial Intelligence : A Modern Approach*, 3rd ed., Prentice Hall, 2010.
- Searle, John R., *Mind: A Brief Introduction*, Oxford University Press, 2004.
- _____, *The Rediscovery of the Mind*, MIT Press, 1994.
- Sebeok, Thomas A. (ed. et al.), *Semiotica*, Journal of the International Association for Semiotic Studies, 2001 Vol. 134(1/4), Mouton de Gruyter Berlin New York.
- Sidel, Victor, “The Right to Health Care: An International Perspective”, in *Bioethics and Human Rights*, ed. by Elsie L. Bandman and Bertram Bandman, New York, 1978.
- Turing, Alan D., “Computing Machinery and Intelligence”, in *Mind*, vol. LIX, No. 236, 1950.
- Wallach, W. · Allen, C., *Moral Machines*, Oxford University Press, 2009.
- URL = https://en.wikipedia.org/wiki/Artificial_intelligence
- URL = <http://platum.kr/archives/6985>

철학탐구 제52집

URL = http://biz.chosun.com/site/data/html_dir/2016/03/20/2016032000637.html

URL = <https://techcrunch.com/2012/01/10/doctors-or-algorithms>

URL = <https://www.youtube.com/watch?v=TTcCkzrjNLY&feature>

Artificial Intelligence and Robot Doctor Ethics

Jooman Maeng (Chung-Ang Univ.)

The existing professional ethics of doctors consists of ethical principles which are centred in the vocation and obligation of doctors, and the relationship between doctors and patients. In the near future robot doctor with on-board the medical artificial intelligence will play important roles of diagnosis and treatment of disease. As the human doctor's reliance on robot doctors increases with this, more and more ethical problems will arise. In that case the human doctor-centred medical ethics limits to solving these problems. Because of this I think we need to enact robot doctor's ethics. It is even more so when considering the expected cooperative relationship in the inevitable times of robot doctor's help in medical practice. This paper is a preliminary consideration for this. I would suggest the fundamental principles of possible medical ethics that should be applied the future robot doctor under the name the ethical principles of robot doctor.

Key words: robot doctor, robot doctor ethics, human doctor, medical ethics, principle of autonomy and mutual consent, principle of nonmaleficence and collective responsibility, principle of beneficence and care, principle of right to medical treatment and justice, human doctor-centred medical ethics, rule- and principle-based moral orientation, ethicbot doctor

철학탐구 제52집

맹주만 E-mail: maeng@cau.ac.kr

투 고 일	2018년 10월 27일
심 사 일	2018년 11월 06일
게재확정	2018년 11월 19일