**빅데이터**

어떤 목적을 가지고 분석을 통해 가치 있는 것으로 창출해낼 수 있는 빠른 속도로 증가하는 대규모의 다양한 데이터(문서, 사진 영상 등)를 이르는 말

**바이오헬스**

생명공학,의·약학 지식에 기초하여 사람이나 다른 생명체에서 유래된 물질을 원료나 재료로 이용하는 기술이다. 특히 유전자 재조합 기술을 이용하는 바이오 의약품, 질병예방 진단 치료가능한 정밀의료, 손상된 조직 장기대체하거나 재생시켜 기능 복원하고 줄기세포를 이용한 바이오장기 관련 산업인 재생의료영역, 헬스케어, 디지털헬스케어 등이 있다.

**바이오헬스 빅데이터**

바이오 데이터 - 생물정보유전체, 생물종정보를 통해 미래 시대 맞춤형 의료서비스를 개선하는데 효율적인 역할을 기대할 수 있다.

헬스 데이터 - 의료정보 및 건강정보를 의미하고 의료기관에서 환자를 진료하면서 생긴 진료정보데이터, 보험 청구를 위해 필요한 보험청구데이터, 건강관련 연구 데이터 등이 있다.

**유전체 정보 빅데이터**

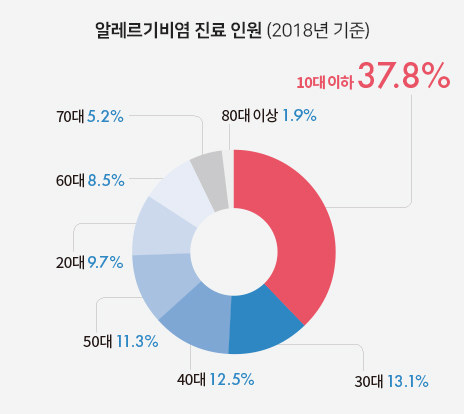
유전체란 우리 몸을 구성하고 있는 생명유지에 필수적인 유전물질이다. 유전체의 기본단위는 DNA인데 사람마다 30억개의 DNA를 가지고 있다. 이 30억개의 DNA에는 개인의 외모적 특성, 질병 민강성에 밀접하게 연관되어 있는 유전 지표라는 것이 있고, 사람마다 약 300만개 이상의 유전지표를 보유한다. 한 개인이 평생 만들어내는 유전지표는 약 6TB 정도이기 유전체 빅데이터라 부른다.

빅데이터 분석으로 개개인이 어떤 질병에 취약하고 발병 가능성은 어느 정도인지 분석 및 예측하고 바이오헬스를 접목하여 이에 따른 개인 정밀 진료 및 건강관리 서비스 등을 적용할 수 있다.

1. 만성질환 환자들이 스스로 질병 관리가 가능하게 POCT 기기를 이용해 자신의 질환을어디서든지 관리할 수 있다.
2. 맞춤형 치료 유전체 진료 임상 data를 이용한다면 환자 개개인에 맞게 분석하여 의료서비스를 제공 하여 진료의 정확도와 치료 효과를 높일 수 있다.
3. 제약회사들이 유전체 정보를 토대로 신약개발을 하면 유전체 데이터를 적용하기 전보다 임상 성공률이 2배 이상 높아지고 처음부터 화합물을 만들지 않고도 치료제를 효율적으로 만들 수 있다.
4. 유전체 빅테이터 분석을 통해 기존약물에 대한 새로운 용도를 밝혀내어 다양한 질병에 사용할 수 있게 된다면 미래에는 약으로 치료가능한 질병이 많아질 것 같다.

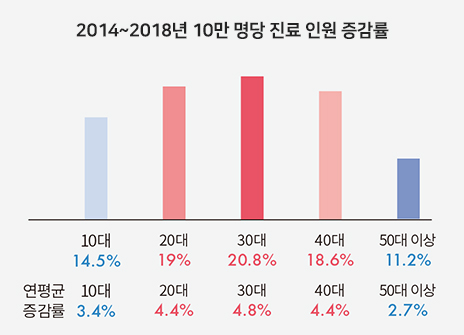
**유전체/바이오헬스 빅데이터 분석을 통해 나이대별 주요질병을 파악한 사례**

**1. 10대 이하 환자가 가장 많은 질환 - 알레르기비염 (사촌동생이 속한 나이 예시)**



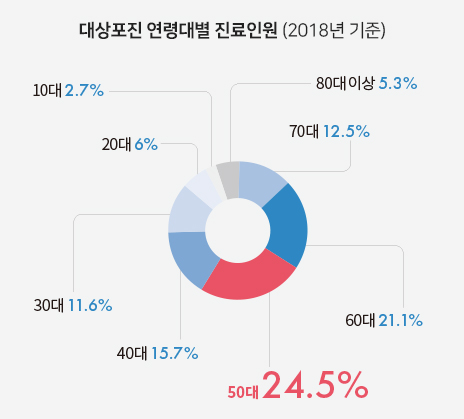
- 알레르기 질환(아토피성피부염, 알레르기비염, 기관지천식)은 순차적으로 발병하고, 나이가 들어감에 따라 증상이 약해진다. 또한 성인보다 편도 아데노이드 비대, 불완전한 부비동의 발달과 부비동염 등의 원인 인자로 증상이 쉽게 나타나기 때문에 10대 이하에서 환자 수가 많은 것으로 보인다.

**2. 20대~40대 환자 증가율이 높은 이석증 ( 대학생 나이)**

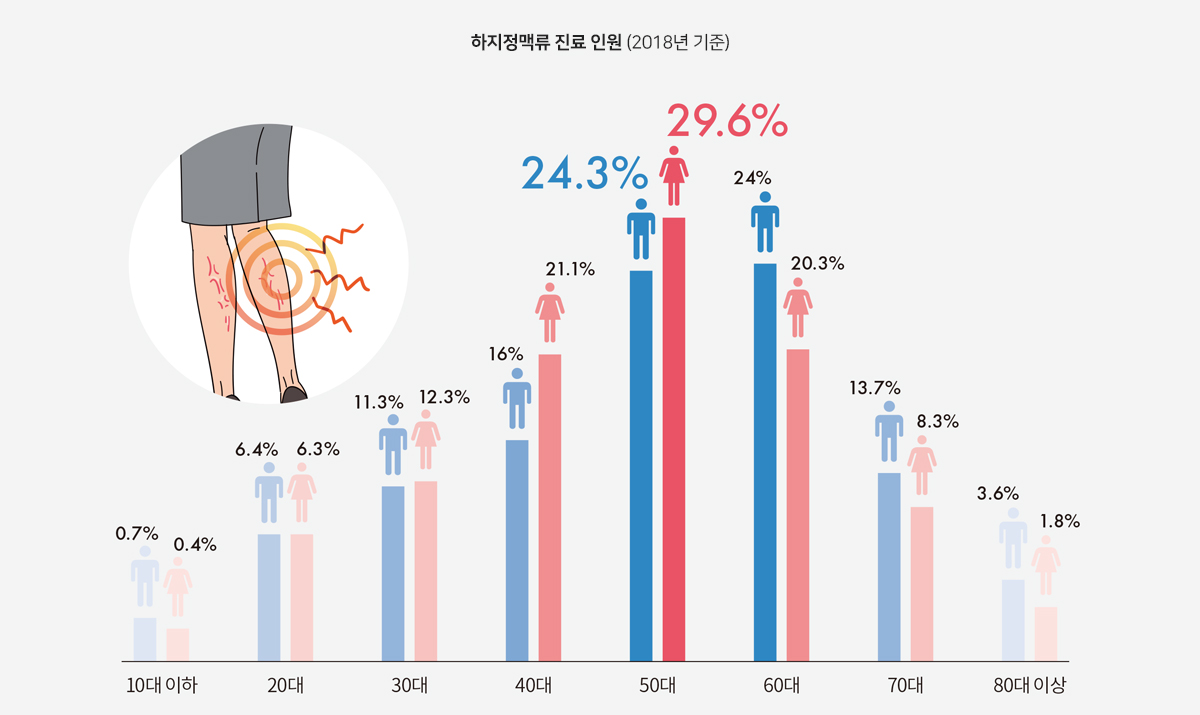


- 5년간 연령대별·성별 10만 명당 진료 인원을 보면 20~40대의 연평균 증감률은 4~5%대로 높게 나타났다. 이는 직업 및 여가 활동이 다른 연령대보다 많은 편이라 두부외상으로 인한 이석증 발생률이 높은 것으로 추정한다.

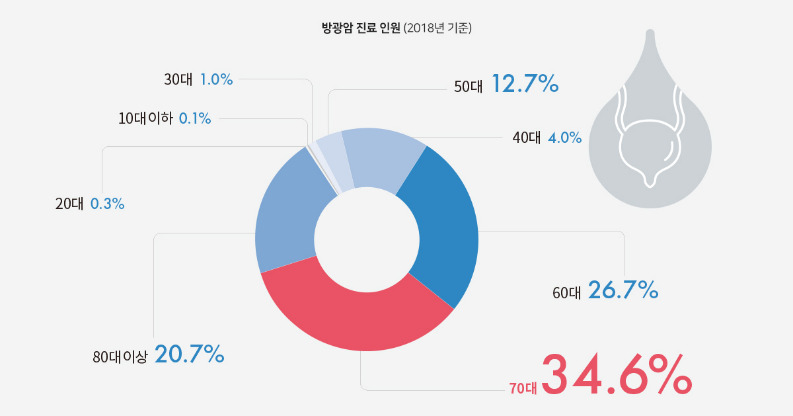
**3. 50대에 많이 나타나는 질환 (부모님 나이 ) --- 대상포진, 하지정맥류**



- 대상포진은 50대 이상 환자가 많은 원인이 확실히 밝혀진 것은 아니다. 하지만 연령이 증가함에 따라 체력이 저하되고, 암이나 당뇨병 같은 만성질환 환자가 많아져 대상포진 환자도 같이 증가하는 것으로 추정한다.



- 하지정맥류는 연령대별로 살펴보면 남녀 모두 50대에서 진료 환자 수가 가장 많았다. 정맥류 발생 원인은 유전적 요인, 직업적 요인 등 여러 가지다. 오래 앉아 있거나 서서 일하는 직업의 경우 시간이 흐를수록 정맥의 탄력이 약화되면서 정맥류가 진행되고, 합병증이 발생할 수 있기에 50대에 병원을 많이 찾는 것으로 보인다. 또 경제적 여건, 직장 문제, 활동 능력 등을 고려할 때 병원 접근성이 가장 용이한 연령층이라는 것도 이유가 될 수 있다.

텍스트이(가) 표시된 사진

자동 생성된 설명**4. 70대 남성이 주의해야할 질병(할아버지 나이) - 방광암**

- 2018년 방광암의 연령대별 진료 현황을 살펴보면 70대 환자(1만2868명, 34.6%)가 가장 많았고, 그다음 60대(9935명, 26.7%), 80대 이상(7694명, 20.7%) 순으로 나타났다. 성별로 보면 남성은 전체 환자의 81%로 대부분을 차지하며, 여성은 19%로 4.2배 차이를 보인다. 남녀 환자 수에 많은 차이를 보이는 것은 암을 유발하는 물질에 대한 생리적 반응의 남녀 차이가 원인일 수 있다. 남성과 여성 빈도의 차이보다 암을 일찍 발견해 치료하는 것이 더 중요하다.

* 국민건강보험공단에서 국민 건강 정보 데이터베이스를 활용해 만든 연령대별 빅데이터는 내가 속한 나이에서 어떤 질병 예방하기 위해 어떻게 조심해야하는지 알 수 있으며, 주요 증상이 나타날 때 어떠한 질환인지 쉽게 예측할 수 있다. 사촌 동생에게도 해당 자료를 보여주며 해당 나이대의 유의해야할 질병과 예방법, 원인 등을 알려주면 빅데이터가 어떻게 활용되는지 더 실감나게 알 수 있을 것이다.
* 또한 요즘 누구나 사용하고 있는 (사촌동생도 사용하고 있는) 스마트폰이나 스마트 워치 같은 일상생활에서 사용하는 건강 관련된 기기들로도 데이터가 수집된다는 사실을 알려주고, 이를 활용하면 환자 개개인에 맞는 치료를 할 수 있어 더 효율적으로 건강을 관리할 수 있게 될 것이라는 것도 알려주면 흥미와 이해를 동시에 자아낼 수 있을 것이다.
* 고령화사회가 되어가면서 미래에는 바이오헬스 빅데이터를 이용한 기술이 점점 많아지고 더욱 일상생활에 밀접한 관계가 생길 것이라 예상한다.

**보건 빅데이터 수집 시 유의할 점**

1. **빅데이터 통계가 믿을 만하고 의지할 만하다는 신념을 줄 수 있어야 한다.**

같은 데이터라도어떻게 가공하냐에 따라 혹은 어떤 알고리즘으로 논리를 만들고 통계를 만들어 내느냐에 따라 매우 다른 결과와 효과를 낸다. 따라서 데이터와 통계자료는 객관적이고 다른 사람들이 쉽게 이해할 수 있도록 간결한 것이 좋습니다.

인공지능 기술이나 바이오헬스 기술력이 뒷받침되어야 실질적이고 효과적인 이득을 안겨줄 수 있으며 사람들이 지속적으로 보건 빅데이터 기술 산업에 투자하고 이를 이용하고자 한다.

1. **정보 보안을 매우 중요하게 고려해야 한다.**

다양한 주체들이 우리의 의료정보를 열람하고 다루게 되는게 이때 동등한 수준의 보안이 유지 되어야 한다. 보안유지가 깨질 경우 이를 악용하거나 부당한 이득을 취하려는 집단이 발생할 수 있기 때문에 유의해야 한다. 특히 유전 정보는 가족의 정보일 뿐만 아니라 가족의 정보이기도 하기 때문에 유전체 빅데이터와 바이오헬스를 다룰 때는 항상 정보 보안 유지에 각별한 주의를 기울여야 사람들이 신뢰할 만하고 긍정적인 이득을 주는 기술로 자리매김할 수 있을 것이다.