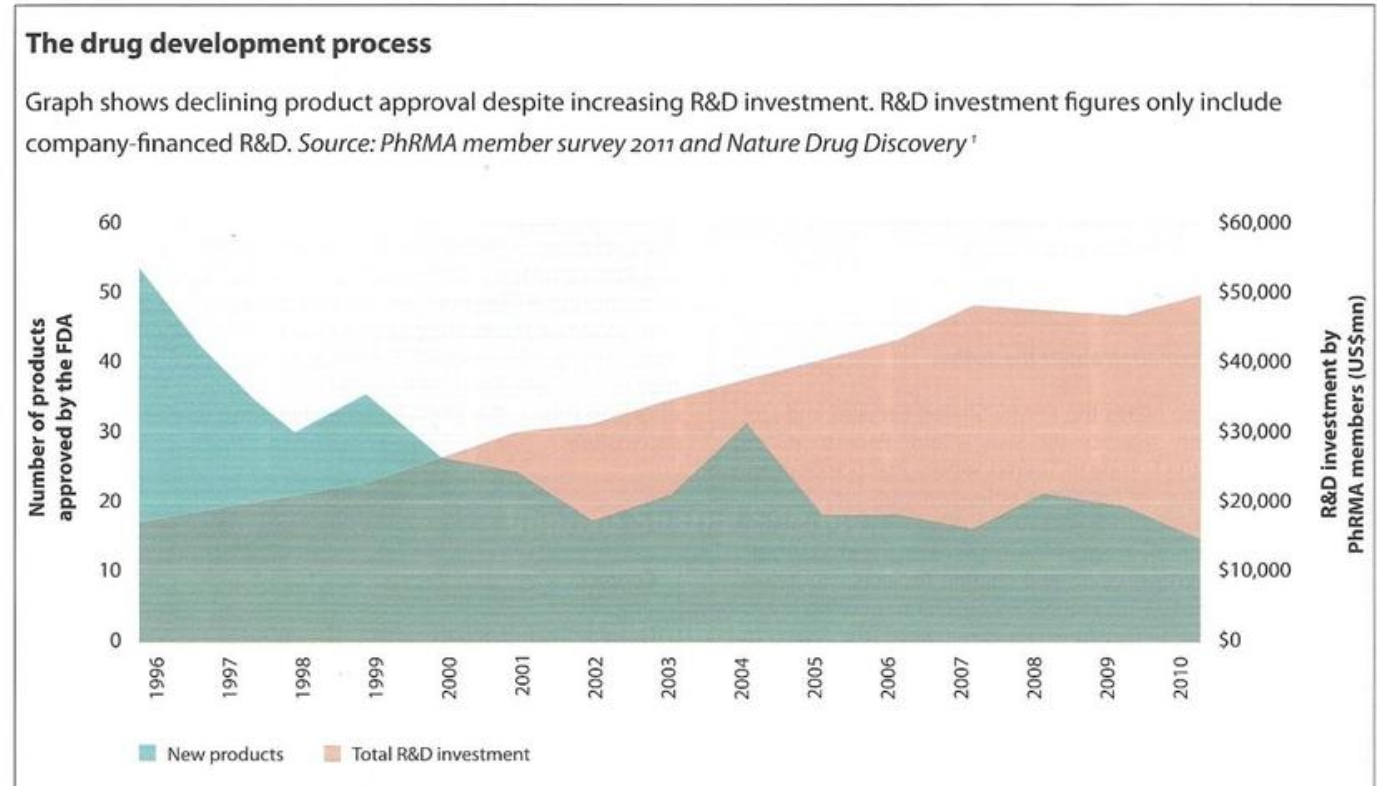
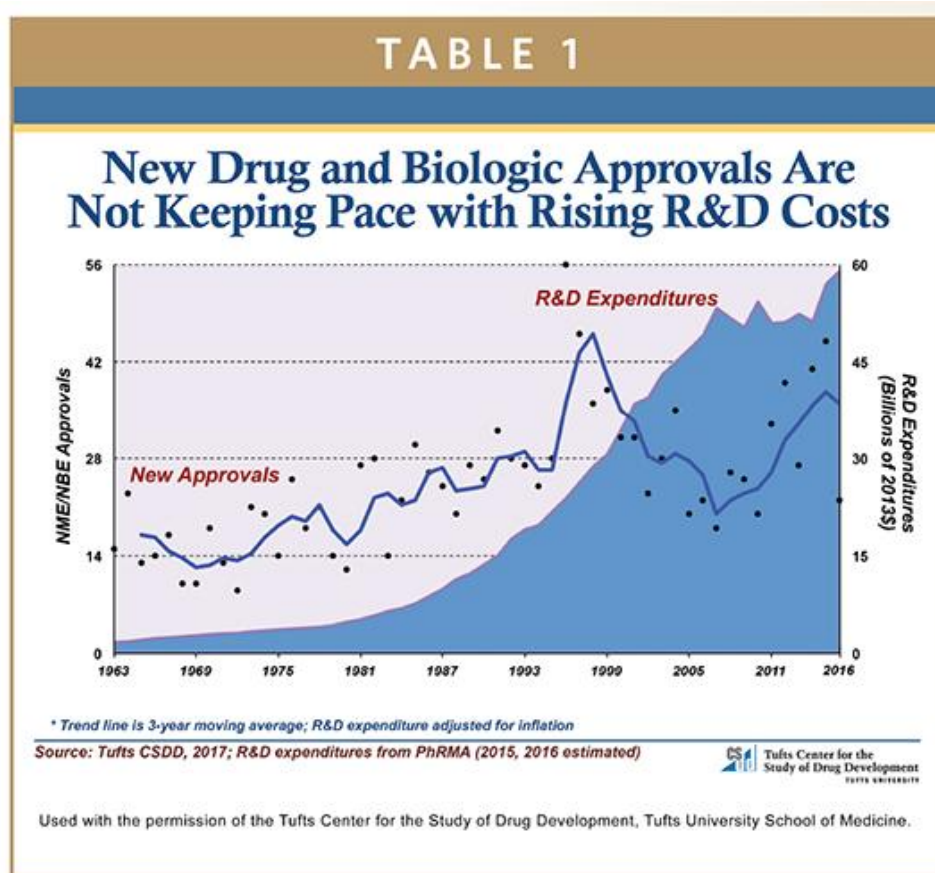
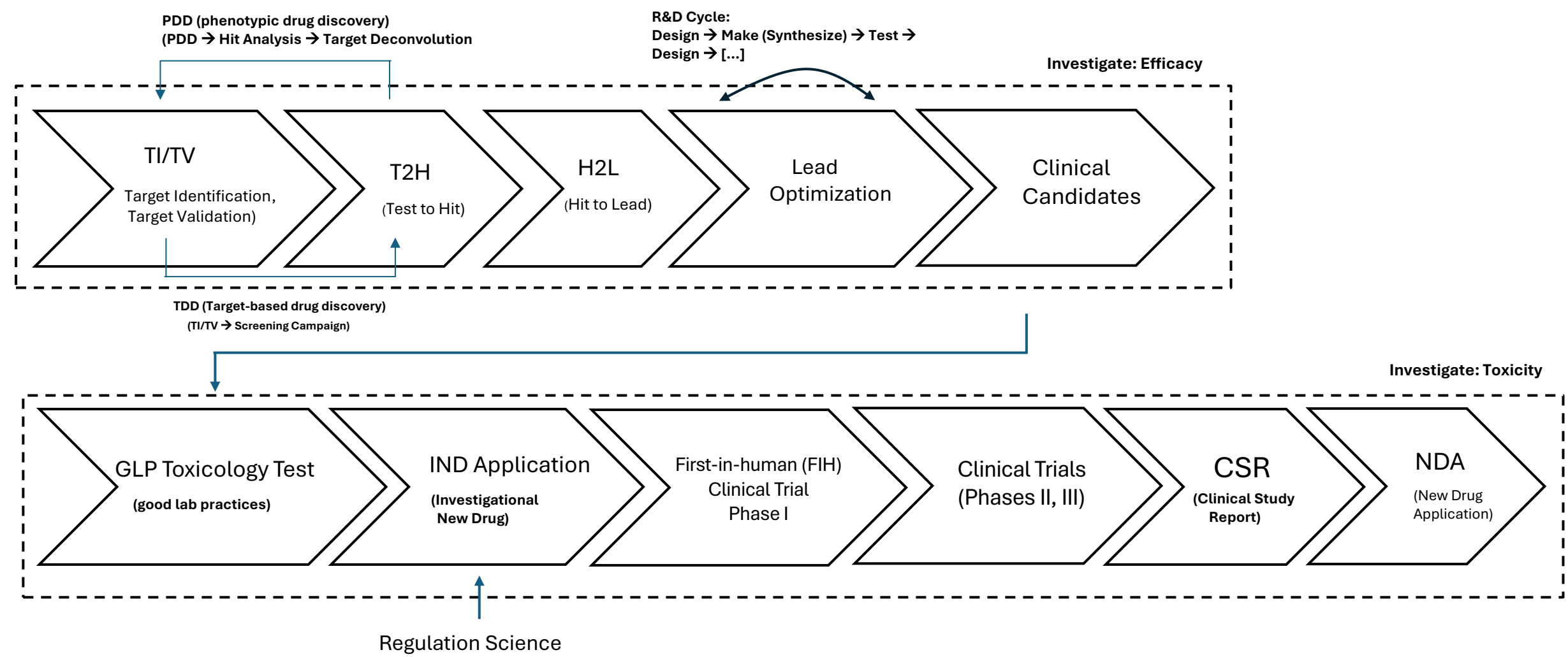


1. Current limitations: 바이오 분야의 '연구개발 패러독스', 즉, 연구를 하면 할수록 약 가격이 천문학적으로 늘고, 또한 새로운 신약 승인 건수가 항상 증가한다는 보장이 없습니다.



¹ Mullard, A, *Nature Reviews Drug Discovery* 10, 82-85 (February 2011)

보통 신약 개발은 임상 3상을 거친 후, 허가를 받는 과정까지 대략 17년 정도가 걸리며, 또한 1만 개 후보 중에 성공하는 것은 3개 정도에 불과합니다.



또한, 신약개발에는 대략 5조 정도의 비용이 들어가며, 이와 같은 높은 연구개발, 생산, 및 유통 비용은 환자의 높은 의료비용에 포함이 됩니다.

High cost of treating cancer

Since 1963, the cost of treating cancer has risen continuously, reaching \$72.1 billion in 2004.

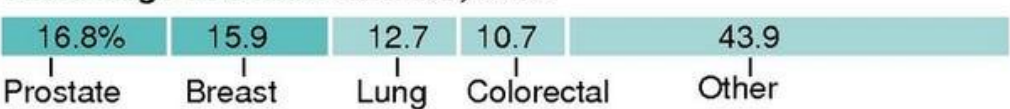
Cancer treatment spending, in billions



Average Medicare payments*, per individual



Percentage of all new cancers, 1998



*In first year following diagnosis, in 2004 dollars

SOURCE: National Cancer Institute

THE ASSOCIATED PRESS

Article | August 30, 2017

Novartis Sets a Price of \$475,000 for CAR T-Cell Therapy

Author(s): [Tony Hagen](#)



Novartis' just-approved chimeric antigen receptor (CAR) T-cell therapy tisagenlecleucel is going to be introduced on the market at a price of \$475,000 for a single infusion, an amount that is within the range anticipated by oncologists.

CAR T-cell therapies have [advanced treatment](#) for patients with certain cancers. They offer sustained remission, fewer side effects, and a short treatment duration. However, they come with a hefty price tag. Between the cost of drug acquisition and administration, as well as management of adverse events (AEs), the cost for Carvykti is [more than half a million dollars](#) per patient per treatment. In some patients, the treatment might need to be repeated.



<https://www.onclive.com/view/novartis-sets-a-price-of-475000-for-car-tcell-therapy>

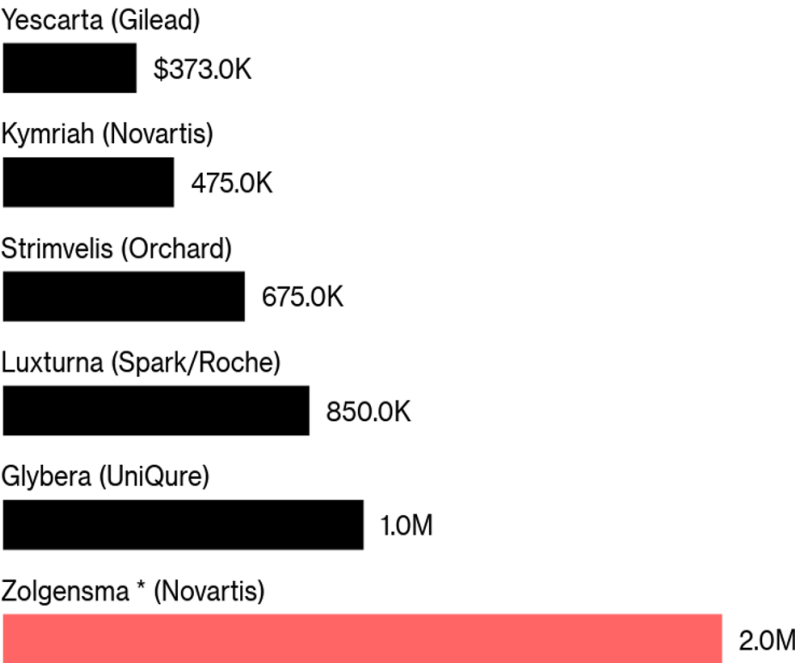
<https://www.managedhealthcareexecutive.com/view/fda-approves-carvykti-for-earlier-treatment-in-multiple-myeloma>

특히 희귀성 유전질환의 경우, 미국 FDA의 승인을 받은 혈우병 치료제 등 유전자 치료제는 가격이 수십 억에 달합니다.

Pricing a Cure

The question for health systems: how much are one-time therapies worth?

■ Price in U.S. dollars



Zolgensma price is UBS assumption and hasn't been decided
Source: Bloomberg

Bloomberg

CNN HealthLife, But BetterFitnessMore

WatchListenLive TV

FDA approves \$3.5 million treatment for hemophilia, now the most expensive drug in the world

(CNN) — A new gene therapy for the fatal genetic disorder metachromatic leukodystrophy, or MLD, will carry a wholesale price of \$4.25 million, its manufacturer announced Wednesday, making it the world’s most expensive medicine.

nature

Explore content>About the journal>Publish with usSubscribe

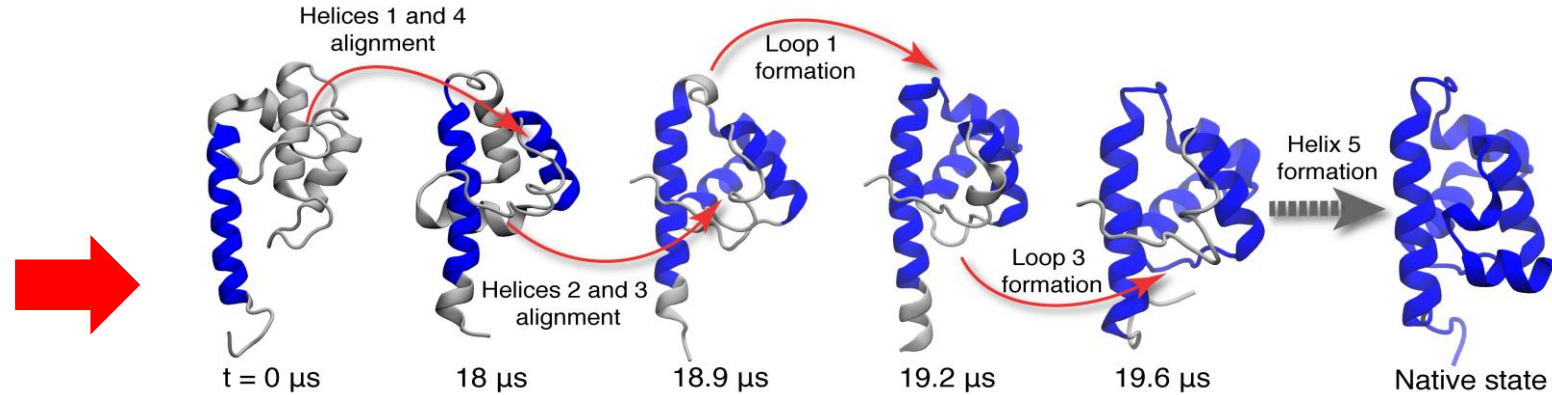
EDITORIAL | 04 December 2019

Expensive treatments for genetic disorders are arriving. But who should foot the bill?

The majority of people with sickle-cell disease are live in the world’s poorest communities and cannot afford the eye-watering costs of treatments.

제약 R&D 과정에서 사용되는 컴퓨터 시뮬레이션은, 여전히 여러 단점들이 있습니다. 예를 들어서, 슈퍼컴퓨터를 사용해도, 단백질 분자 동역학(MD) 시뮬레이션은 시스템의 복잡성으로 인하여 계산만 몇 달이 걸릴 수 있습니다.

supercomputer



~ microsecond simulation = ~ months (at best) computing time

MD simulations, which track the motions of every atom in a large molecule, are limited in their time scales by computer power and architecture. It can take months for a supercomputing system to simulate only tens of microseconds of a protein's dynamics. But many proteins fold on the millisecond time scale.

Borman, Stu. "Anton Supercomputer Proves Mettle." *Chemical & Engineering News*, vol. 88, no. 42, 18 Oct. 2010, cen.acs.org/articles/88/i42/Anton-Supercomputer-Proves-Mettle.html. Accessed 23 Feb. 2025.

또한, 슈퍼컴퓨터에서 시뮬레이션을 실행하는 데는 전기, 냉각, 유지 관리 등 상당한 운영 비용이 발생하며, 이는 곧 환경 오염 및 에너지 고갈 등의 문제점으로 이어집니다.

Carnegie Mellon University

Mellon College of Science

About

News & Events

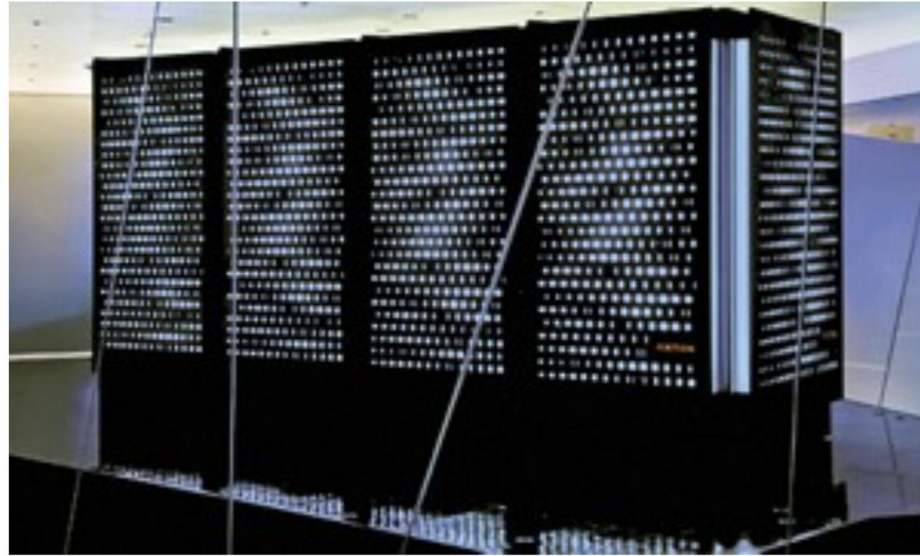
Undergraduate

Graduate

August 06, 2024

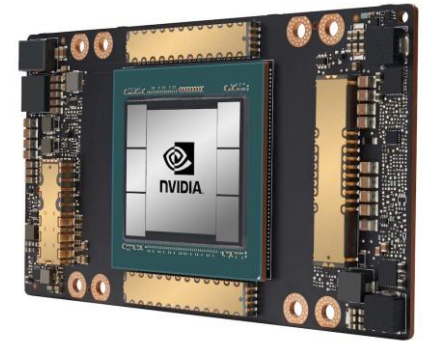
\$3.15 Million from NIH to Fund Operation of Third-Generation Anton Supercomputer at Pittsburgh Supercomputing Center

Designed and Built by D. E. Shaw Research, System Will
Simulate Biomolecules Roughly 100 Times Faster than
General-Purpose Supercomputers



Credit: Courtesy of Matthew Monteith

*Shaw Research's supercomputer Anton is
designed to simulate the dynamics of biological
macromolecules.*



NVIDIA A100 GPU : USD \$25,000

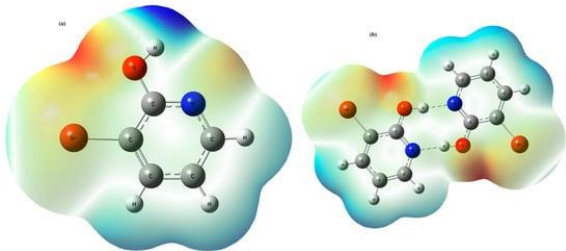
GPU X 1000 = \$ 25,000,000

Chiacchia, Kenneth. "\$3.15 Million from NIH to Fund Operation of Third-Generation Anton Supercomputer at Pittsburgh Supercomputing Center." *Mellon College of Science, Carnegie Mellon University*, 6 Aug. 2024, www.cmu.edu/mcs/news-events/2024/0806_3-15-million-from-nih-to-fund-operation-of-third-generation-anton-supercomputer.html. Accessed 23 Feb. 2025.

2. Quantum Solution: 양자 컴퓨터는 중첩, 얽힘, 간섭의 원리를 통하여, 여러 변수를 동시에 처리하는 병렬 처리(parallel processing) 방법 등으로, 단백질 시뮬레이션 등 복잡한 문제에 대해 훨씬 **자연스럽고**, 빠르게 계산을 수행할 수 있습니다.



Molecular Simulation

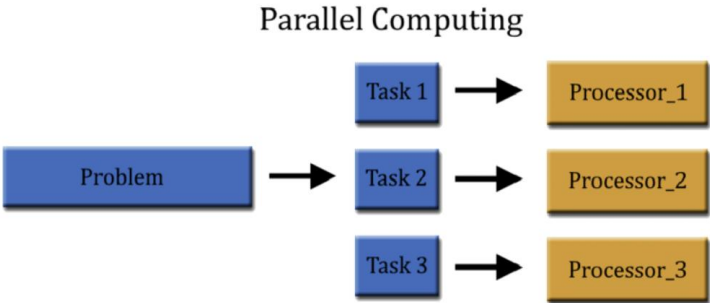
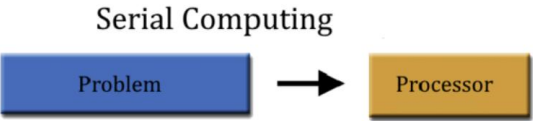


Quantum Mechanics

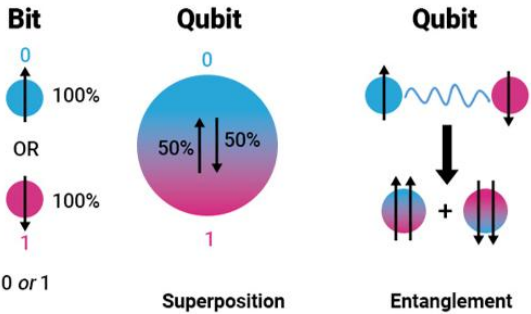
$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi$$

Natural
Language

Information Unit: Qubit

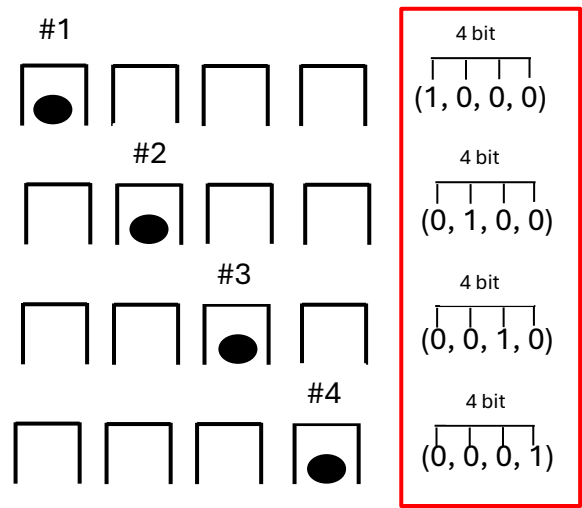


Natural
Information
Processing



상자 속 동전을 찾는 간단한 예제로, 양자 알고리즘의 속도 측면 이점을 확인해볼 수 있습니다.

Classic Approach: Sequential Search



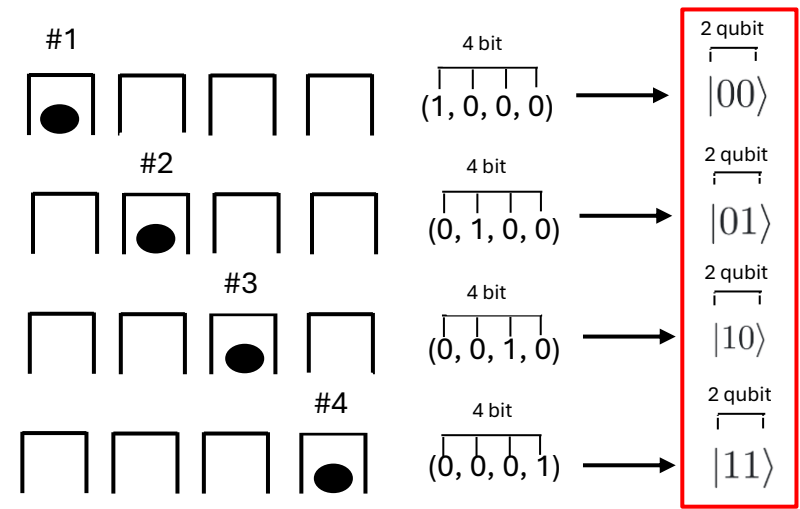
Task: Search Box with Coin → Python Code

```
def find_coin(boxes):  
    for i in range(4): # 4개의 상자를 순서대로 확인  
        if boxes[i] == 1:  
            return i + 1 # 상자 번호 (1부터 시작)  
    return -1 # 동전이 없는 경우 (이 문제에서는 발생하지 않음)  
  
boxes = [0, 1, 0, 0]  
print(find_coin(boxes)) # 출력: 2 (두 번째 상자)
```

총 필요한 순차적 iteration 횟수 (즉, ‘상자 확인 변수’) = 대략 N번. [~ O(N) order]
→ (e.g) 상자가 4개면, 한 번에 한 개의 상자를 각각 뒤집는 순차적인 확인을, 총 4번 해야 함.

상자 속 동전을 찾는 간단한 예제로, 양자 알고리즘의 속도 측면 이점을 확인해볼 수 있습니다.

Classic Approach: Sequential Search



Task: Search Box with Coin → Python Code

```
def find_coin(boxes):  
    for i in range(4): # 4개의 상자를 순서대로 확인  
        if boxes[i] == 1:  
            return i + 1 # 상자 번호 (1부터 시작)  
    return -1 # 동전이 없는 경우 (이 문제에서는 발생하지 않음)  
  
boxes = [0, 1, 0, 0]  
print(find_coin(boxes)) # 출력: 2 (두 번째 상자)
```

총 필요한 순차적 iteration 횟수 (즉, ‘상자 확인 변수’) = 대략 N번.
→ (e.g) 상자가 4개면, 한 번에 한 개의 상자를 각각 뒤집는 순차적인 확인을, 총 4번 해야 함.

Quantum Approach: Grover’s Algorithm (1996)

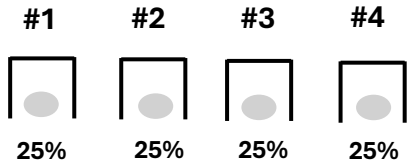
Task: 상자 밑에 동전이 있는 상자를 찾으시오. (정답: 2번 상자)

$|\psi\rangle$: 각 상자 밑 동전의 존재 확률을 표현한 현재 양자 상태

출발 조건

$$|\psi\rangle = \frac{1}{2} (|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

측정
양자 상태의 붕괴



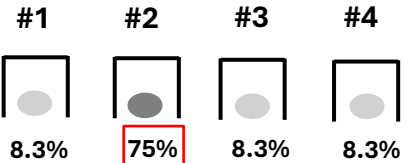
Iteration #1

- 1. Oracle 연산 via CNOT/Z gate, etc
- 2. 확률 증폭 (amplitude amplification)

(e.g.)

$$|\psi\rangle = \frac{1}{\sqrt{12}} (|00\rangle + 3|01\rangle + |10\rangle + |11\rangle)$$

측정
양자 상태의 붕괴



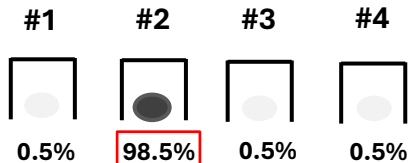
Iteration #2

- 1. Oracle 연산자
- 2. 확률 증폭

(e.g.)

$$|\psi\rangle = \frac{1}{\sqrt{103}} (|00\rangle + 10|01\rangle + |10\rangle + |11\rangle)$$

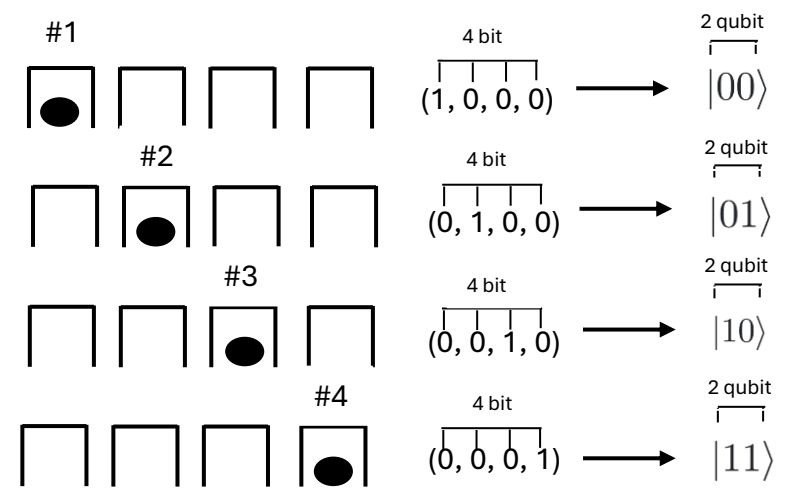
측정
양자 상태의 붕괴



ψ 를 측정할 시에, ψ 붕괴 후 98.5% 확률로 2번 상자를 선택함

상자 속 동전을 찾는 간단한 예제로, 양자 알고리즘의 속도 측면 이점을 확인해볼 수 있습니다.

Classic Approach: Sequential Search



Task: Search Box with Coin → Python Code

```
def find_coin(boxes):  
    for i in range(4): # 4개의 상자를 순서대로 확인  
        if boxes[i] == 1:  
            return i + 1 # 상자 번호 (1부터 시작)  
    return -1 # 동전이 없는 경우 (이 문제에서는 발생하지 않음)  
  
boxes = [0, 1, 0, 0]  
print(find_coin(boxes)) # 출력: 2 (두 번째 상자)
```

총 필요한 순차적 iteration 횟수 (즉, ‘상자 확인 변수’) = 대략 N번.
→ (e.g) 상자가 4개면, 한 번에 한 개의 상자를 각각 뒤집는 순차적인 확인을, 총 4번 해야 함.

Quantum Approach: Grover’s Algorithm (1996)

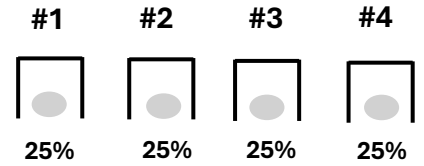
Task: 상자 밑에 동전이 있는 상자를 찾으시오. (정답: 2번 상자)

$|\psi\rangle$: 각 상자 밑 동전의 존재 확률을 표현한 현재 양자 상태

출발 조건

$$|\psi\rangle = \frac{1}{2} (|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

측정
양자 상태의 붕괴



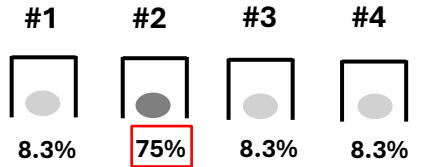
Iteration #1

- 1. Oracle 연산 via CNOT/Z gate, etc
- 2. 확률 증폭 (amplitude amplification)

(e.g.)

$$|\psi\rangle = \frac{1}{\sqrt{12}} (|00\rangle + 3|01\rangle + |10\rangle + |11\rangle)$$

측정
양자 상태의 붕괴



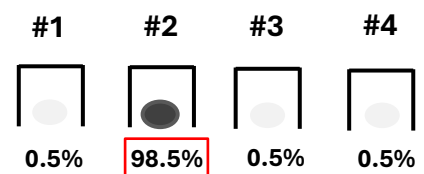
Iteration #2

- 1. Oracle 연산자
- 2. 확률 증폭

(e.g.)

$$|\psi\rangle = \frac{1}{\sqrt{103}} (|00\rangle + 10|01\rangle + |10\rangle + |11\rangle)$$

측정
양자 상태의 붕괴



ψ를 측정할 시에, ψ 붕괴 후 98.5% 확률로 2번 상자를 선택함

총 필요한 순차적 iteration 횟수 = 대략 Root(n)번. [~ O(√n) order]

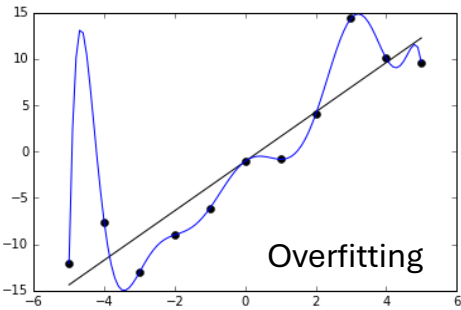
→ (e.g) 상자가 4개면, 한 번에 모든 상자를 뒤집는 확인 (‘oracle 연산’)을, 총 root(4) = 2번만 하면 됨.
→ 이와 같은 고전적 계산 방식 대비 이점을 ‘Quadratic Speedup’이라 함.

의료 데이터는 multimodal 구조, 변수 간 collinearity 가능성, unknown 변수 및 high-dimensionality 특성 등을 포함하여 매우 복잡합니다. 이러한 복잡한 데이터를 처리하고 유의미한 정보를 추출하는 데에서, 고전적 통계 알고리즘은 여러 한계가 있습니다.

문제 1. High dimensional data

- 정의: 변수 개수 >> 관측점 개수인 상황
- 문제점:
 - Overfitting
 - 차원의 저주 (Curse of Dimensionality)
 - Non-deterministic model

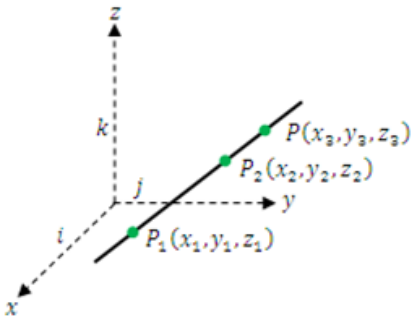
예시: (a,b) 2차원 점 11개를 계수 $n \gg 11$ 개인 n 차 함수로 fitting하는 경우



문제 2. Collinearity

- 정의: 많은 독립변수들이 같은 방향으로 동시에 증가/감소하는 상황
- 문제점:
 - 모수 추종 단계서, 모수들의 variance가 커지고, 모델이 불안정해짐.
 - 귀무 가설 기각이 어려워짐으로, 통계적 검증이 무의미해짐

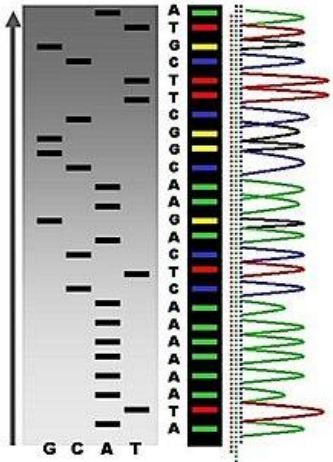
예시: 3차원 공간에서 x,y,z가 같은 방향으로 증가하는 모습



문제 3: Multimodality

- 정의: 이질적 (heterogeneous)인 데이터 유형 (type)이 존재하는 상황
- 예시: 전자 의료 기록 (EMR), 유전체 데이터, 의료 이미지/영상 데이터
- 문제점: 복잡한 cross-modality 관계 및 패턴을 어떻게 capture할 수 있는지 등, 난제 존재.

.txt



.jpg



mp4



예를 들어서, 공선성(collinearity)이 존재할 경우, 회귀계수 추정의 불안정화, 표준오차의 증가, t-검정값의 감소 등으로 인하여, 개별 변수의 통계적 유의성을 판단하기 어려워집니다.

Cancer Risk = $\beta_0 + \beta_1$ (Smoking) High correlation
+ β_2 (Alcohol Consumption) + ϵ

→ Find $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ with least squares method

Linear 2-variable Regression: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

1. Data

Y	X ₁	X ₂
2	1	1
4	2	0
6	3	1
8	4	0
10	5	1

2. Design matrix (X)

$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & 4 & 0 \\ 1 & 5 & 1 \end{bmatrix}$

3. Compute: $X^T X, X^T Y, (X^T X)^{-1}$

Compute: $\hat{\beta} = (X^T X)^{-1} X^T Y$

$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$

4. Compute: $Var(\hat{\beta}), Var(\hat{\beta}_1), Var(\hat{\beta}_2)$

$Var(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \rightarrow \epsilon \sim N(0, \sigma^2)$
 $= \begin{bmatrix} Var(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2) \\ Cov(\hat{\beta}_1, \hat{\beta}_2) & Var(\hat{\beta}_2) \end{bmatrix}$

높은 값

$Var(\hat{\beta}_1) = \frac{\sigma^2}{(1 - \rho^2) S_{11}}$

$Var(\hat{\beta}_2) = \frac{\sigma^2}{(1 - \rho^2) S_{22}}$

$S_{11} = \sum (X_1 - \bar{X}_1)^2,$
 $S_{22} = \sum (X_2 - \bar{X}_2)^2,$
 $S_{12} = \sum (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$

상관계수 $\rho = \frac{S_{12}}{\sqrt{S_{11} S_{22}}}$ | $\rho \rightarrow 1$
 $1 - \rho^2 \rightarrow 0$, 일수록 $Var(\hat{\beta}_1) \gg 1$
 $Var(\hat{\beta}_2) \gg 1$

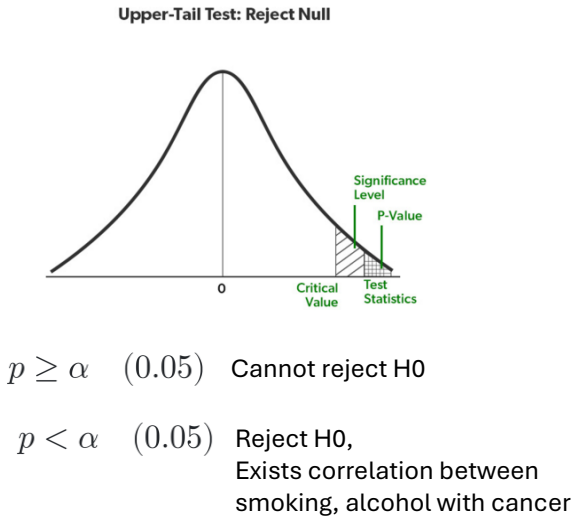
5. Conduct t-test

$H_0 : \beta_1 = 0$ $t_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
 $H_A : \beta_1 > 0$

$H_0 : \beta_2 = 0$ $t_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} = \frac{\hat{\beta}_2}{\sqrt{Var(\hat{\beta}_2)}}$
 $H_A : \beta_2 > 0$

p-value₁ = $P(T > t_1 | H_0)$

p-value₂ = $P(T > t_2 | H_0)$



6. Colinearity가 높은 경우

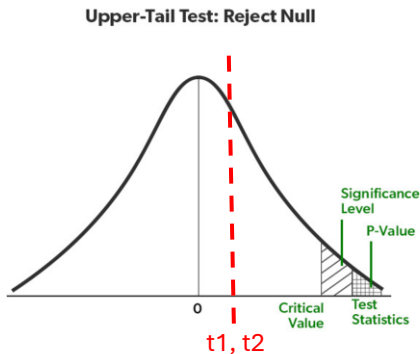
$\rho = \frac{S_{12}}{\sqrt{S_{11} S_{22}}} \rightarrow 1$

$Var(\hat{\beta}_1) \gg 1$

$Var(\hat{\beta}_2) \gg 1$

$t_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}} \ll 1$

$t_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} = \frac{\hat{\beta}_2}{\sqrt{Var(\hat{\beta}_2)}} \ll 1$



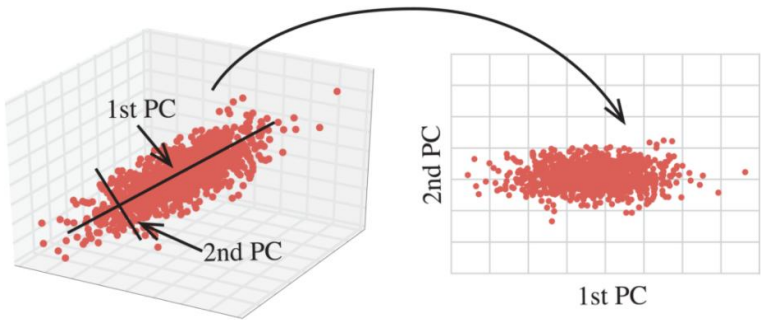
❖ Fail to reject H0

통계적 검정으로 흡연 및 음주와 암 유발률 사이의 상관관계 밝히는데 실패

High dimensional data 및 collinearity가 존재하는 경우, 이를 PCA와 같은 차원축소법 등, 고전적 통계 기법으로 분석해볼 수 있습니다. 다만, 이와 같은 고전적 통계 기법은 데이터의 양이 굉장히 많아지는 경우, 속도 및 효율성에서 한계를 보입니다.

주성분 분석 (PCA)

<PCA 분석>



<PCA 기법 도식 (Step 1~5) >

1. n개 변수를 가진 데이터 행렬 정규화 수행 (즉, $X \rightarrow Z$ 변환) (각 변수 x_i 마다, 평균=0, 표준편차=1로 정규화)
2. 데이터 행렬의 Covariance (공분산) 행렬 C 계산
3. C 의 (λ, v) pair 계산 $\rightarrow (\lambda_1, v_1), (\lambda_2, v_2), \dots, (\lambda_m, v_m)$ (단, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$)
단, $C \in \mathbb{R}^{n \times n}$ 에서 $n \gg 1$ 인 경우, **QR algorithm**을 사용해서 $(\lambda_1, v_1), (\lambda_2, v_2), \dots, (\lambda_m, v_m)$ 계산
4. 주성분 행렬 계산: $Z_{PCA} = Z \times V = [Z_1, Z_2, \dots, Z_n]_{n \times m} \times [v_1, v_2, \dots, v_m]_{m \times m}$, $Z_{PCA} \in \mathbb{R}^{n \times m}$
 $\rightarrow Z_{PCA} = [PC_1, PC_2, \dots, PC_m]_{n \times m}$
5. 생성한 principal component (PC) 중, 고유값 크기 순으로 k개 PC 선택 $\rightarrow Z_{PCA} = [PC_1, PC_2, \dots, PC_k]_{n \times k}$

<**QR algorithm** 도식> 차원 $n \times n$ 인 공분산 행렬에서 eigenvalue 추출 과정: $O(n^3)$ 번 계산 복잡도

$$\text{공분산 행렬 } C \longrightarrow C = Q_0 R_0$$

$$C_1 = R_0 Q_0 \longrightarrow C_1 = Q_1 R_1$$

$$C_2 = R_1 Q_1 \longrightarrow [\dots] \quad C_k = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_m \end{bmatrix}$$

\longrightarrow : Gram-Schmidt 직교화

R: Upper triangular 행렬

Q: Orthogonal 행렬 (즉, $Q^T Q = I$)

다만, PCA와 같은 통계적 기법에서, 양자 알고리즘의 속도적 우위가 입증된 연구가 있습니다.

고전적 PCA (주성분분석):

$$[X_1, X_2, \dots, X_n] \quad (n \gg 1) \longrightarrow [PC_1, PC_2, \dots, PC_k] \quad (k \ll n)$$

변수가 n 개인 데이터 행렬의
Principal component 계산 과정:
 $O(n^3)$ 번 계산 복잡도

$$PC_i = a_{1i}X_1 + a_{2i}X_2 + \dots + a_{ni}X_n$$

$$\begin{bmatrix} PC_1 \\ PC_2 \\ \vdots \\ PC_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

※ PC: Principal Components (주성분)

Discussion

Density matrix exponentiation represents a powerful tool for analysing the properties of unknown density matrices. The ability to use n copies of ρ to apply the unitary operator $e^{-i\rho t}$ allows us to exponentiate non-sparse d -dimensional matrices to accuracy $\epsilon = O(t^2/n)$, and to perform qPCA to construct the eigenvectors and eigenvalues of a low-rank matrix ρ in time $O(R \log d)$. Like quantum matrix inversion¹⁴, qPCA maps a classical procedure that takes time polynomial in the dimension of a system to a quantum procedure that takes time polynomial in the logarithm of the dimension.

Quantum Advantage:

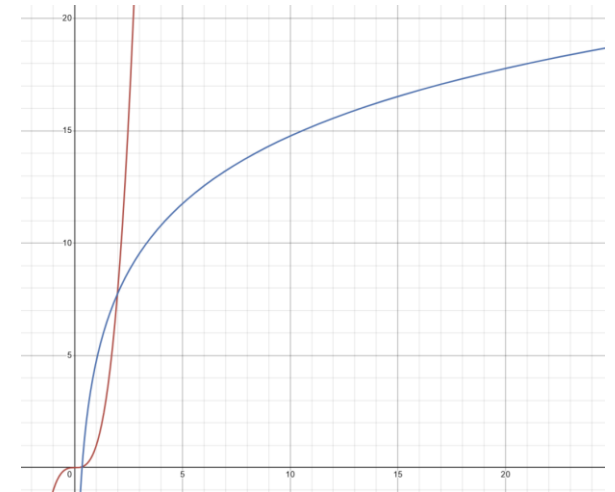
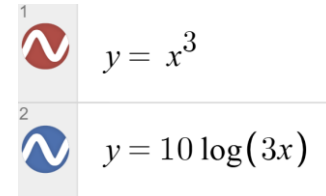
1. Process larger datasets faster
2. Find more complex relationships/patterns in data
3. Better at handling high-dimensional data (PCA)

Quantum PCA 알고리즘:

변수가 m 개인 $n \times m$ 행렬의 주성분 (PC)을 계산할 때,
대략 $O(R \cdot \log(nm))$ 번 계산 필요. (R: 행렬 rank, nm = 행렬 차원)

→ $n \gg 1$ 일 경우, $O(R \cdot \log(nm)) \ll O(n^3)$

→ Principal component 계산 과정에서,
이와 같은 속도 이점을 logarithmic speedup 이라 함.



nature
physics

LETTERS

PUBLISHED ONLINE: 27 JULY 2014 | DOI: 10.1038/NPHYS3029

Quantum principal component analysis

Seth Lloyd^{1,2*}, Masoud Mohseni³ and Patrick Rebentrost²

Lloyd, Seth, Masoud Mohseni, and Patrick Rebentrost. "Quantum Principal Component Analysis." *Nature Physics*, vol. 10, no. 9, 2014, pp. 631-633. doi:10.1038/nphys3029.

예를 들어서, 양자 검색 (search) 및 분류 (clustering) 알고리즘을 활용하면, 암 환자의 유전체 데이터를 분석하여, 병원성 유전자 변이나 암 연관 돌연변이를 더욱 신속하게 식별 및 분류할 수 있습니다.

1. Cancer Diagnosis



2. Multi-omics data

- Genomic Data
- Transcriptomic Data
- Proteomic Data
- Metabolomic Data

Data
Preprocessing

3. Quantum Algorithms

(e.g.) Search

- Genome/DNA sequence
 - Find Pathogenic gene variation compared to normal physical gene variation

- Transcriptomic (scRNA-seq) Sequence
 - Dysregulated/Aberrant pathways

(e.g.) Clustering/Classification

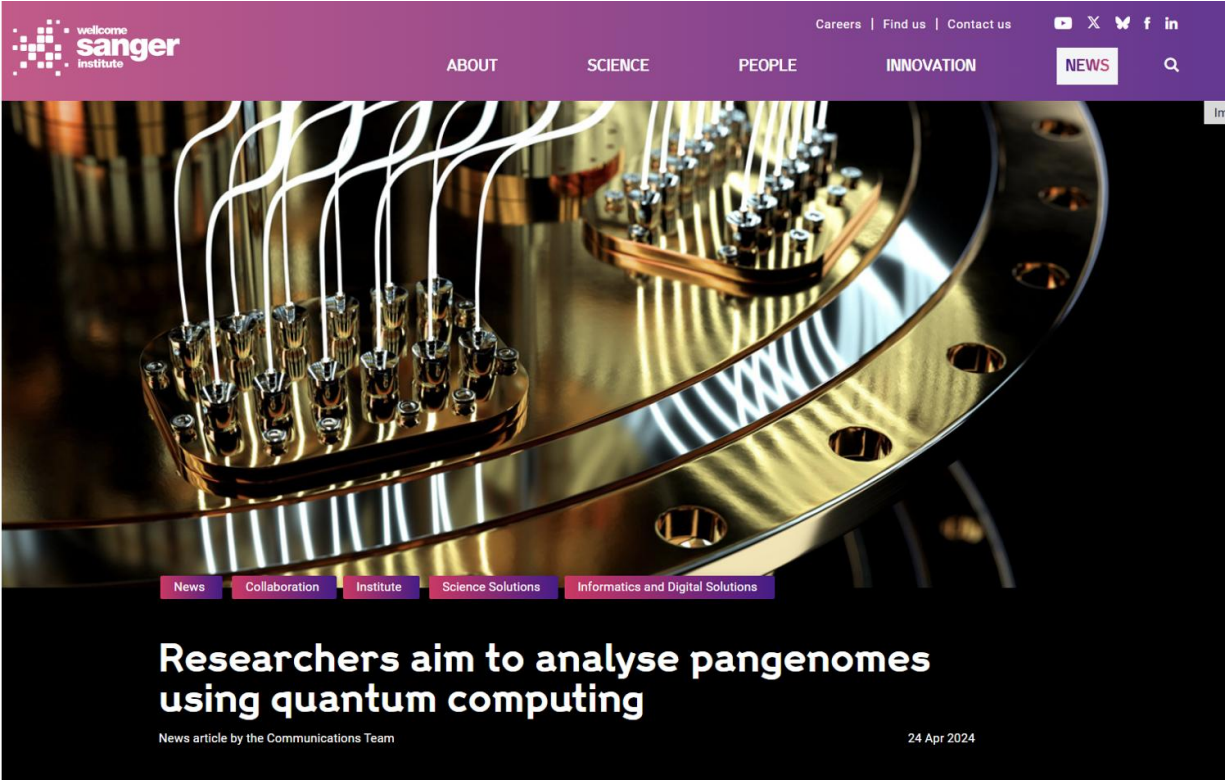
- Cluster tumors into subtypes based on different molecular signatures

- Find correlations (btw genes/pathways) across different omics layers → holistic tumor modeling

4. Personalized medicine

- (e.g.) Tumor metabolic profile
 - metabolic dependency-targeting therapy
- (e.g.) Tumor immune profile
 - tumor-specific immunotherapy

양자 컴퓨팅을 인구 유전체학 연구에 적용한 한 예시: 2024년 케임브리지大 연구진은 양자 컴퓨터로 판게놈 데이터를 구축, 보완 및 분석하는 것을 목표로 최대 350만 달러(약 50억원)의 연구비를 확보했습니다.



- Useful as reference genome in:
 - Disease Treatment
 - Population Genetics / Evolution Studies
 - Anthropology



About EMBL > EMBL Communications > News Related: [Press office](#) [All EMBL sites](#)

Researchers aim to use quantum computing to assemble and analyse pangenomes

Experts in quantum computing and genomics to develop new methods and algorithms to process biological data

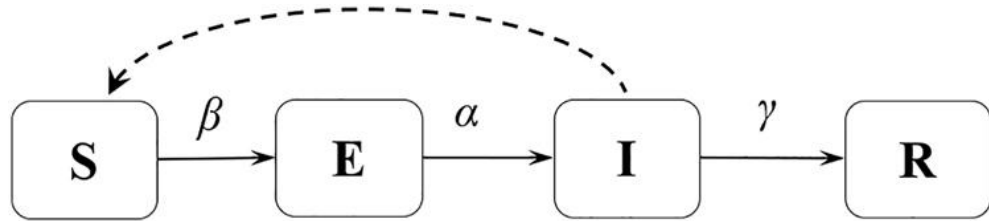


IBM Q System One Quantum Computer at the Consumer Electronic Show in 2020. Credit: AA+W Adobe Stock

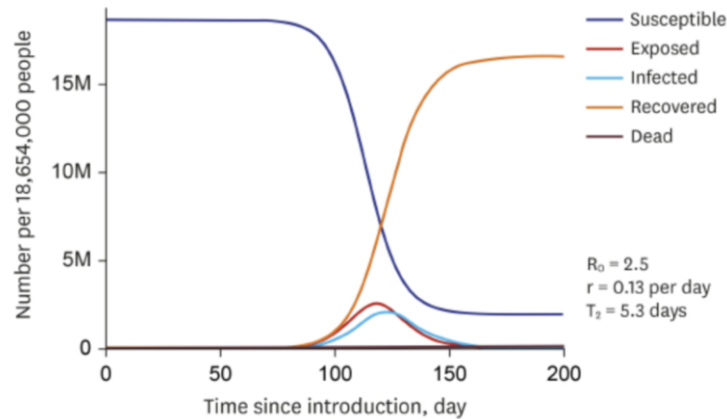


https://www.sanger.ac.uk/news_item/researchers-aim-to-analyse-pangenomes-using-quantum-computing/

또한, 양자 최적화 알고리즘은 코로나-19 등 질병 확산 모델에서 파라미터를 더 정확하게 예측하는데 사용될 수 있으며, 이를 통하여 질병 확산에 대한 효과적인 방역 전략을 수립할 수 있습니다.



$$\begin{aligned}\frac{dS}{dt} &= -\beta IS/N \\ \frac{dE}{dt} &= \beta IS/N - \sigma E \\ \frac{dI}{dt} &= \sigma E - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$



S: Susceptible individuals
 E: Exposed individuals (infected but not yet infectious)
 I: Infectious individuals
 R: Recovered or removed individuals

β (beta): The transmission rate
 γ (gamma): The recovery rate
 σ (sigma): 1/latent period

Input:

- Cost function $C(\beta, \sigma, \gamma)$ to be minimized.

Quantum Optimization \rightarrow Parameter Estimation

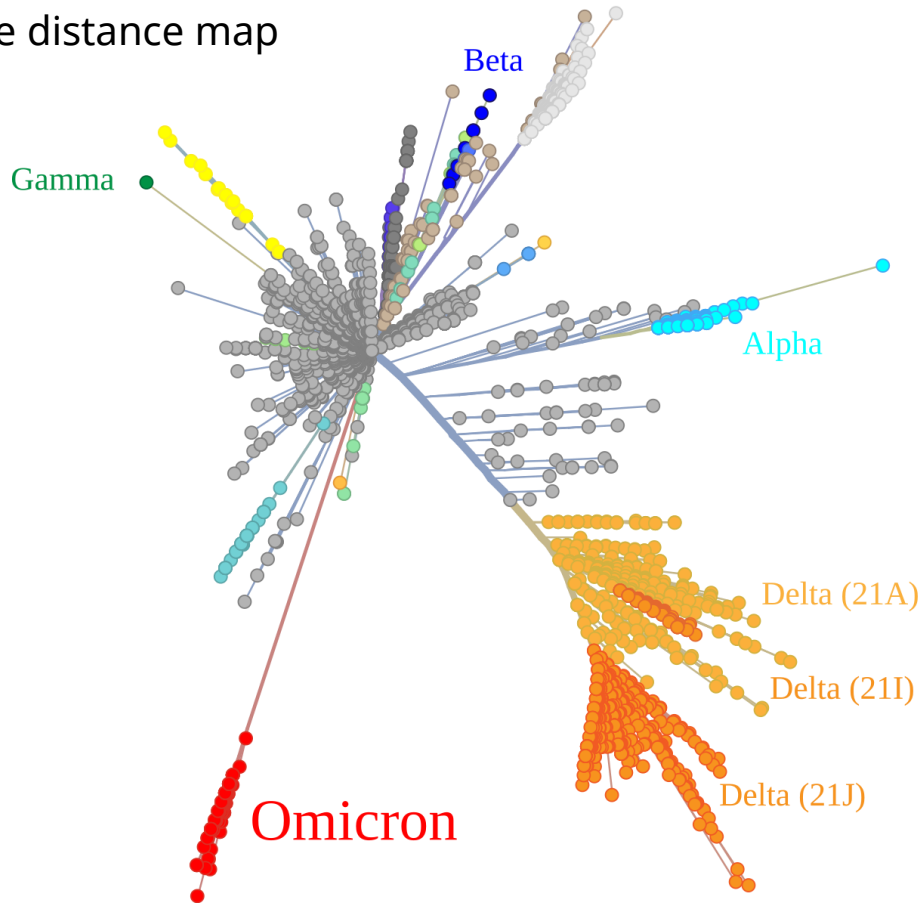
$$\begin{aligned}\theta^{(t+1)} &= \theta^{(t)} - \alpha \nabla C(\theta^{(t)}) \\ \theta &= (\beta, \sigma, \gamma)\end{aligned}$$

Output:

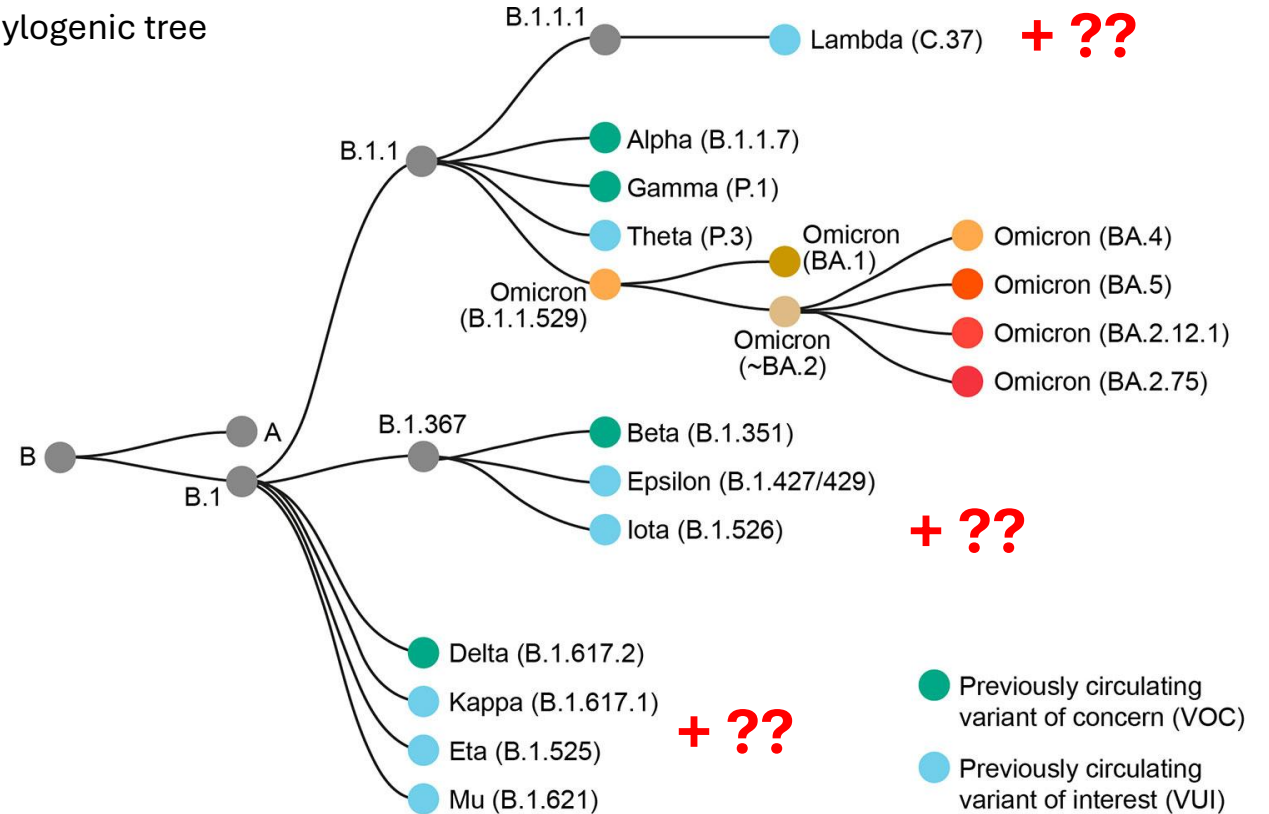
- Optimal parameter set $(\beta^*, \sigma^*, \gamma^*)$.

글로벌 팬데믹 상황에서, 양자 알고리즘은 바이러스 변이체 간의 유전적 거리를 계산하고, 바이러스 계통발생학적 트리를 분석할 수 있습니다. 이를 통해서, 전염성 증가, 또는 면역 회피로 이어질 수 있는 돌연변이를 식별하여, 잠재적으로 위험한 변이체를 예측할 수 있습니다.

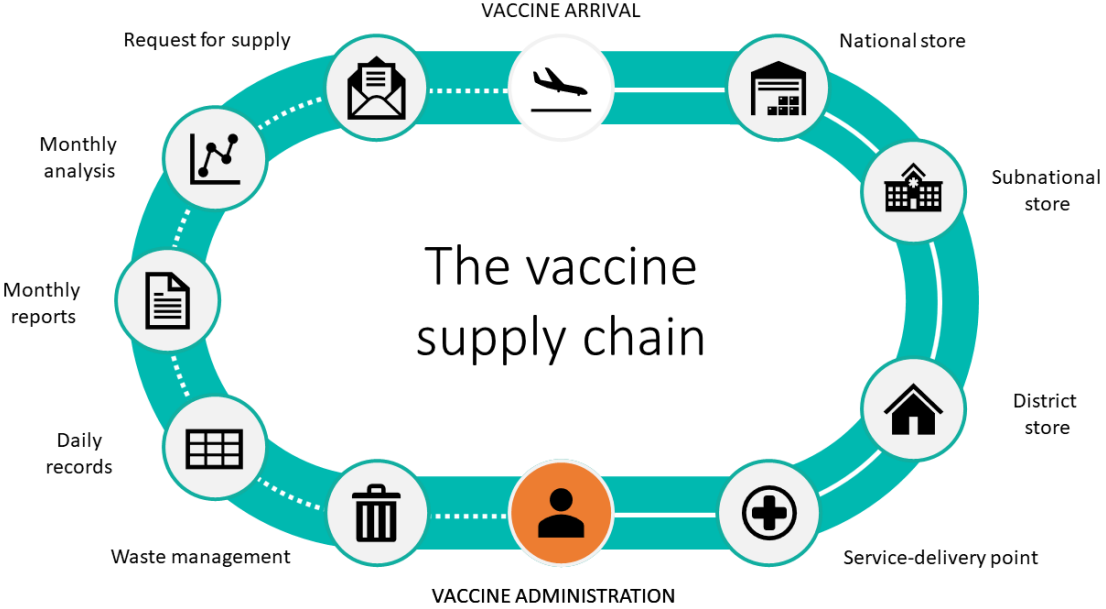
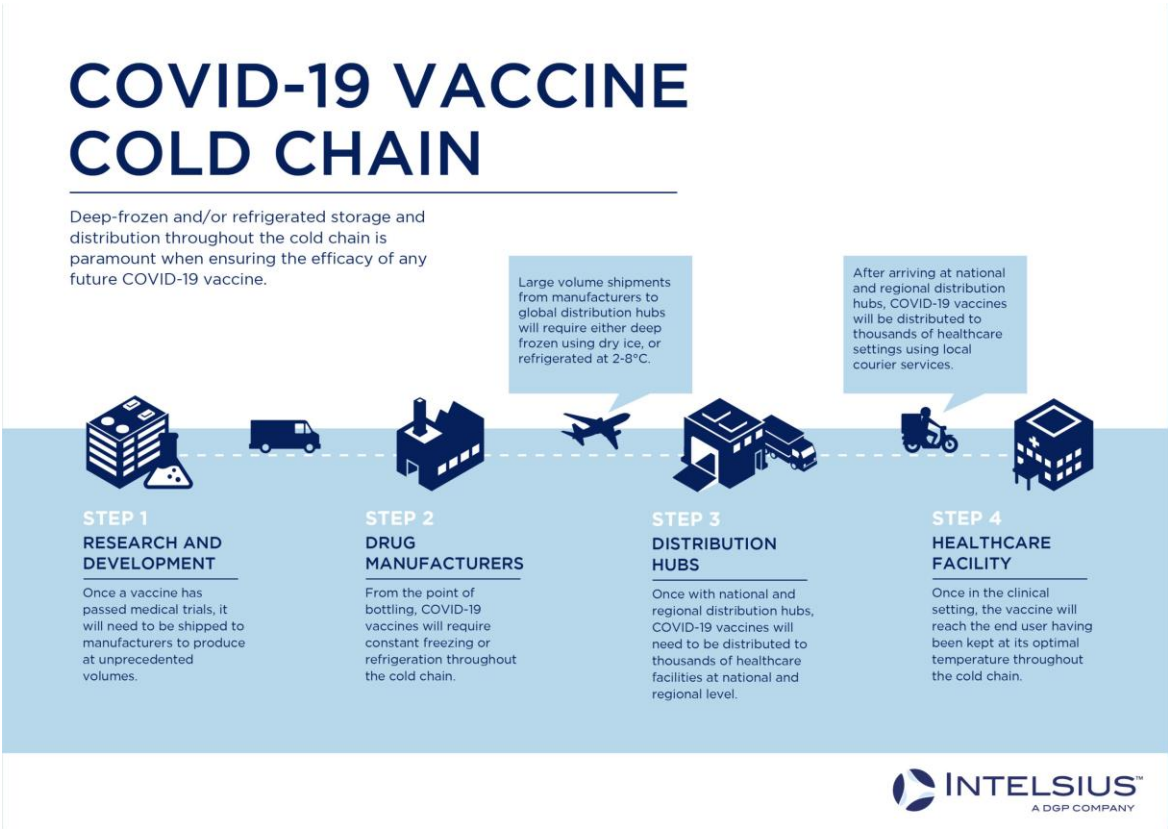
SARS-CoV-2 mutation gene distance map



SARS-CoV-2 variants Phylogenetic tree



또한, 양자 알고리즘을 활용하면 COVID-19와 같은 팬데믹 등 글로벌 보건 위기에서 복잡한 공급망 운영을 최적화하고, 자원 배분의 효율성을 극대화할 수 있습니다.



3. Conclusion. 양자컴퓨터를 활용해 보건 분야의 난제들을 해결하고, 특정 인류가 소외되는 일 없이 공영할 수 있도록 노력이 필요합니다.

THE LANCET Global Health

[This journal](#) [Journals](#) [Publish](#) [Clinical](#) [Global health](#) [Multimedia](#) [Events](#) [About](#)

EDITORIAL · Volume 12, Issue 3, E341, March 2024 · [Open Access](#)

[Download Full Issue](#)

The landscape for rare diseases in 2024

[The Lancet Global Health](#)

[Article Info](#) [Linked Articles \(1\)](#)

[Download PDF](#) [Cite](#) [Share](#) [Set Alert](#) [Get Rights](#) [Reprints](#)



By definition, rare diseases affect a small number of individuals (fewer than 1 in 2000 people in any WHO region); yet, with more than 7000 types of rare disease in existence, the burden worldwide is not insignificant. To date, approximately 300 million people live with rare diseases. Such individuals are often a neglected and marginalised group, especially those in low-income and middle-income countries. Around 80% of **rare diseases**

Saha, Sujata, et al. "Trends in Maternal Mortality, 2000–2020: Estimates by WHO, UNICEF, UNFPA, World Bank Group, and UNDESA/Population Division." The Lancet Global Health, 22 Feb. 2024, [www.thelancet.com/journals/langlo/article/PIIS2214-109X\(24\)00056-1/fulltext](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(24)00056-1/fulltext). Accessed 23 Feb. 2025.