

精準醫學 與 大數據

楊永正 (yang@ym.edu.tw)

國立陽明大學生物醫學資訊研究所

這是一個顛覆傳統的時代

The infographic features a central white vertical pole with four dark grey arrows pointing outwards. Each arrow points to a company name and a descriptive sentence. The companies and their descriptions are: Uber (The world's largest taxi company, owns no vehicles.), Facebook (The world's most popular media owner, creates no content.), Alibaba (The most valuable retailer, has no inventory.), and Airbnb (The world's largest accommodation provider, owns no real estate.).

Uber

The world's largest taxi company, owns no vehicles.

The world's most popular media owner, creates no content.

Facebook

Alibaba

The most valuable retailer, has no inventory.

The world's largest accommodation provider, owns no real estate.

Airbnb

Something interesting is happening.
TOM GOODWIN

wetp@int
creative digital solutions

<https://www.rebelmouse.com/wetpaintmena/>

© / WetpaintMENA

產生破壞式革新的技術： 改變生活、商務、與全球經濟的進展

- Mobile internet
(移動互聯網)
- Automation of knowledge work
- The internet of things
- Cloud technology
- Advanced robotics
- Autonomous and near-autonomous vehicles
- Next generation genomics
- Energy store
- 3D printing
- Advanced materials
- Advanced oil and gas exploration and recovery
- Renewable energy

- by McKinsey Global Institute, May, 2013



台北公車

mywoo - 2013年6月12日
交通運輸

安裝

加入願望清單

★★★★☆ (3,430)

台北公車(市政府) 紅10 09:49:14

公車代號 網站

去程 陽明高中:4分 | 回程 士林國中:8分

台北海專	預計到站時間	小北街	預計到站時間
台北海專	目前未發車	小北街	4分
浮線	目前未發車	士林區農會	6分
富洲里九段	目前未發車	士林國中	8分
	目前	士林區	

愛 群 紅 綠 藍 它

台北公車 群體A 09:49:51

紅10

去程: 4分 陽明高中

回程: 8分 士林國中

橘1

去程: 尚未發車 建業路

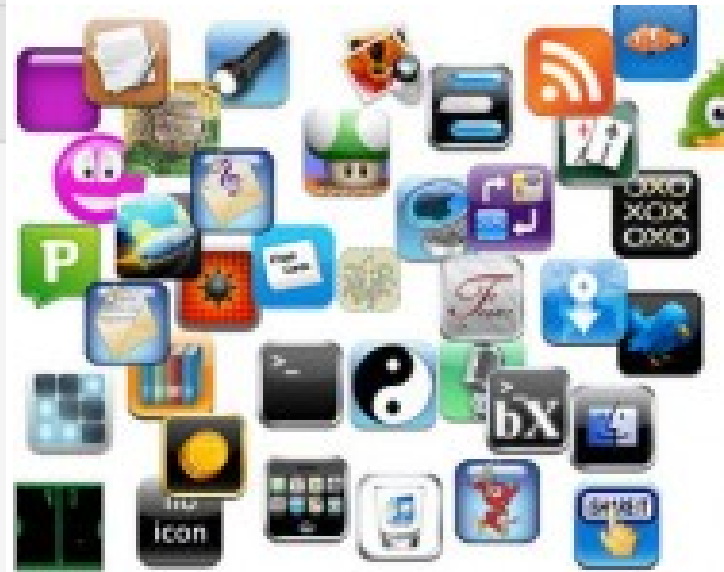
信義新幹線

去程: 3分 世貿中心

回程: 6分 衡陽路口

群體 A 群體 B 群體 C

APP大幅增加生活的方便性



https://play.google.com/store/apps/details?id=com.mywoo.busplus&hl=zh_TW

<http://www.techweekeurope.co.uk/news/amazon-android-appstore-briefly-opens-outside-us-40287>

移動互聯網的成功需要 開放數據(open data)

DATA.GOV.TW
政府資料開放平臺 (公開測試版)

搜尋

首頁

資料分類

使用規範

常見問題

關於我們

分類 機關別

- 生育保健 (3)
- 出生及收養 (4)
- 求學及進修 (35)
- 求職及就業 (14)
- 開創事業 (3)
- 投資理財 (9)
- 休閒旅遊 (43)
- 交通及通訊 (31)

更多

資料類型

- 原始數據
- 行動化應用軟體
- 工具軟體
- 系統介接程式

首頁 > 資料分類 > 最新訊息與熱門資料

NEW 僑務委員會議資料表

CSV

更新日期：民國102年12月01日

瀏覽人次：8

NEW 高雄港郵輪船期

CSV

更新日期：民國102年9月12日

瀏覽人次：1155

NEW 楠梓.高雄園區就業資訊

CSV

更新日期：民國102年9月11日

瀏覽人次：3566

NEW 車型耗能證明核發資料

CSV

更新日期：民國102年9月10日

瀏覽人次：923

NEW 國際環保新聞週報

CSV JSON XML

更新日期：民國102年9月09日

瀏覽人次：3

HOT iTaiwan中央行政機關室內公共區域免費無線上網熱點查詢服務

CSV

更新日期：民國102年4月26日

瀏覽人次：14898

HOT 電影

EXCEL JSON XML

更新日期：民國102年5月03日

瀏覽人次：7339

HOT 展覽資訊

EXCEL JSON XML

更新日期：民國102年5月02日

瀏覽人次：4051

政府機構有明確的開放政策，並提供使用說明，可擴大民間參與的機會

回首頁

關於我們

資料目錄

APPs

開發指南

最新消息

常見問答

使用規範

» 現在位置 [首頁](#) > 開發指南

資料搜尋：



開發指南

友善列印

Data.Taipei 臺北市資料開放平台提供各式資料，民眾或機關可在網站下載或以API介面方式取得資料，俾供後續加值及應用。

為方便開發人員取得及使用，本平台以較普及的檔案格式和業界一般的標準提供資料。

提供方式：

- 資料下載：直接於網站中的各資料頁面，依據瀏覽器使用方式點選該檔案以滑鼠右鍵下載
檔案格式：Excel、CSV、MDB、XML、KML、SHP、DGN等
- 系統API 介接：可直接以系統介接的方式連結該資料介接URL取得內容。
檔案格式：JSON

每份資料均有詮釋資料，列出資料的格式及欄位定義，協助開發人員充分了解資料內容或應使用的介接明文件。

» [提供API介接之資料項目清單](#)：

想像若能串接
公車時間與電
影開演時間，
是否讓你更容
易規劃時程？

開放數據的品質影響佈署之方便性

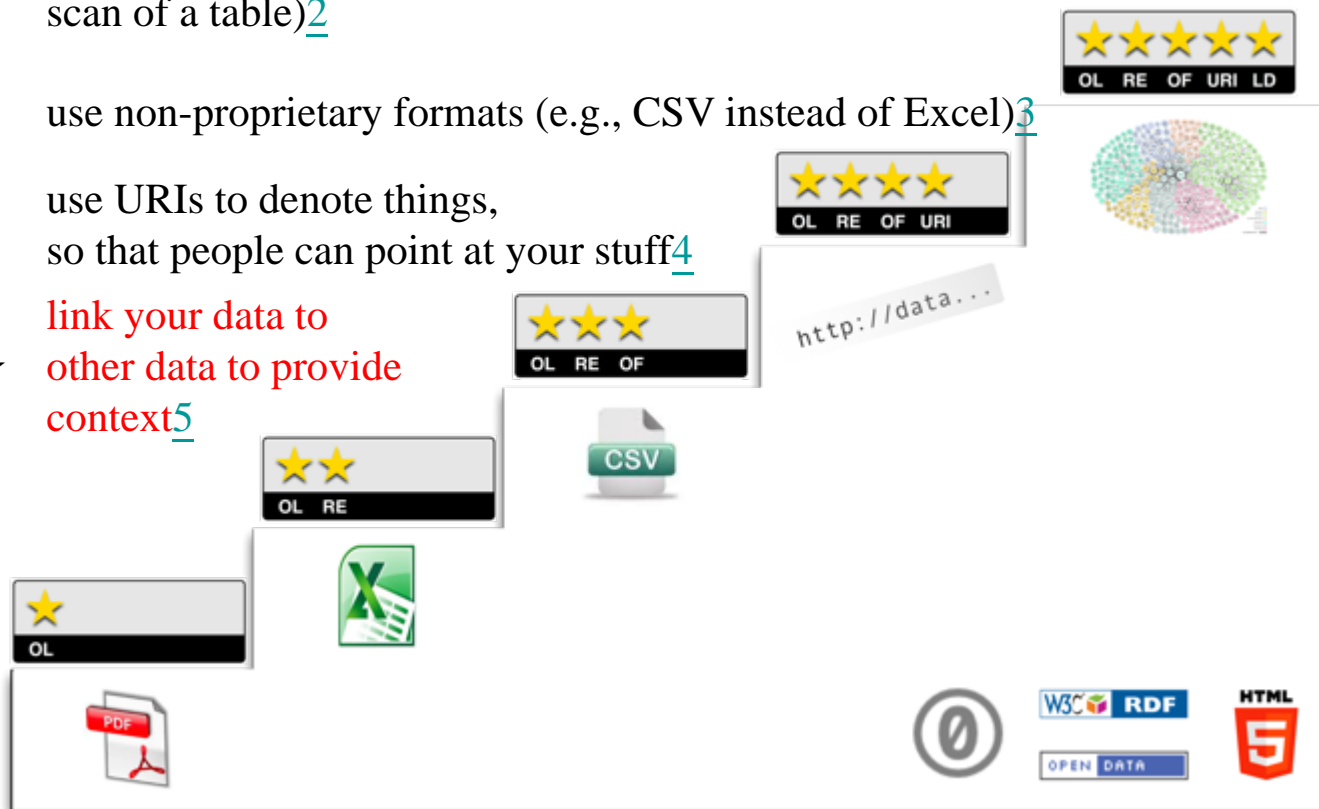
★ make your stuff available on the Web (whatever format) under an open license¹

★★ make it available as structured data (e.g., Excel instead of image scan of a table)²

★★★ use non-proprietary formats (e.g., CSV instead of Excel)³

★★★★ use URIs to denote things, so that people can point at your stuff⁴

★★★★★ link your data to other data to provide context⁵



4G (LTE)網路是否能取代wifi尚不清楚，
但「無線上網」已是生活的一部份



Sri Lanka will be the first country in the world to have universal Internet coverage.

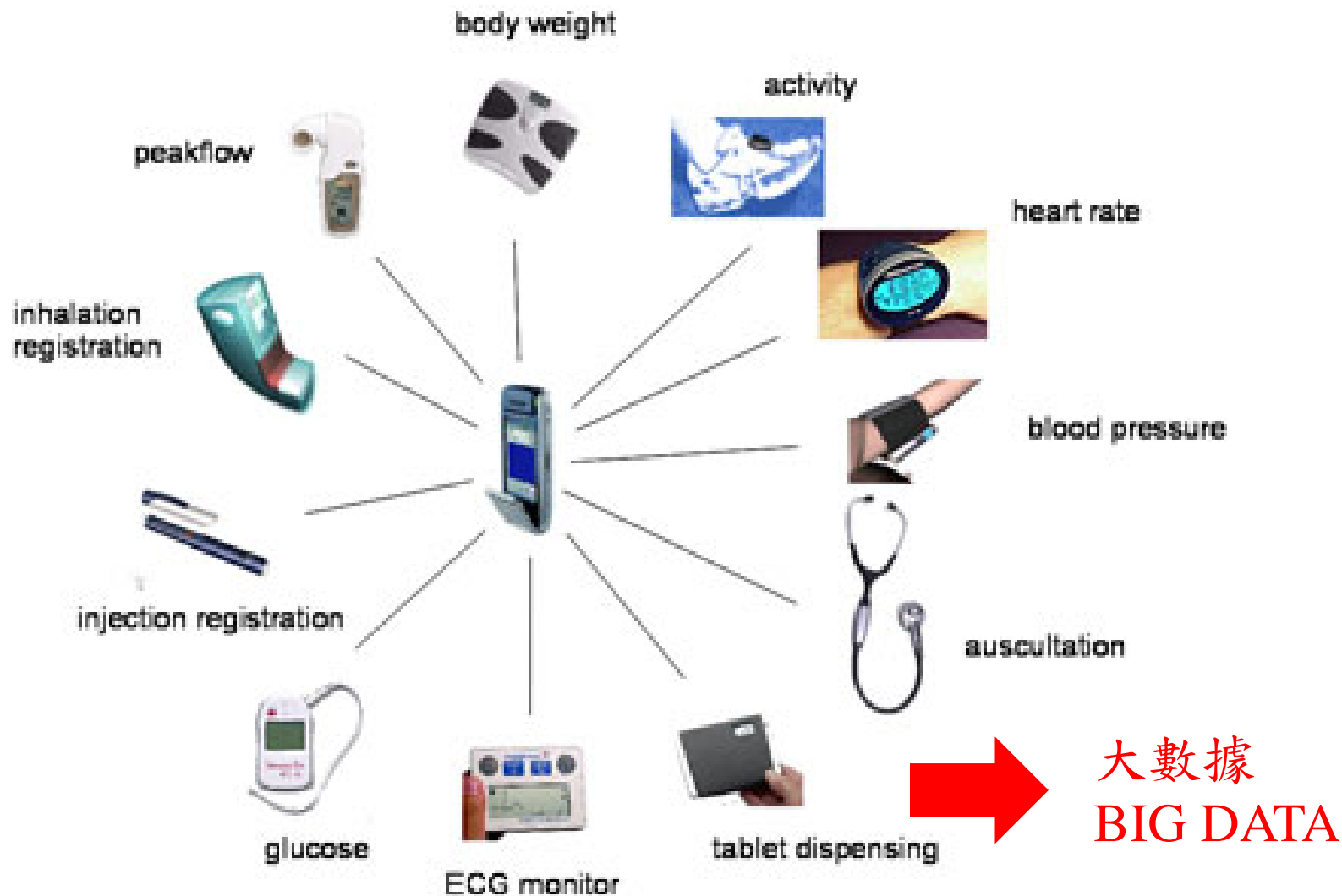
<http://www.sciencealert.com/google-s-internet-balloons-will-soon-connect-all-of-sri-lanka-with-wi-fi>

當無線上網成為生活的一部份，萬物皆可上網，物聯網儼然成型

- Mobile internet
- Automation of knowledge work
- The internet of things (物聯網)
- Cloud technology
- Advanced robotics
- Autonomous and near-autonomous vehicles
- Next generation genomics
- Energy store
- 3D printing
- Advanced materials
- Advanced oil and gas exploration and recovery
- Renewable energy

- by McKinsey Global Institute, May, 2013

物聯網: 越來越多的設備可以透過 網路收集各種有用的數據



當數據大到與母群體差不多，桌機的計算力
比以前的伺服器更大，還需要取樣嗎？

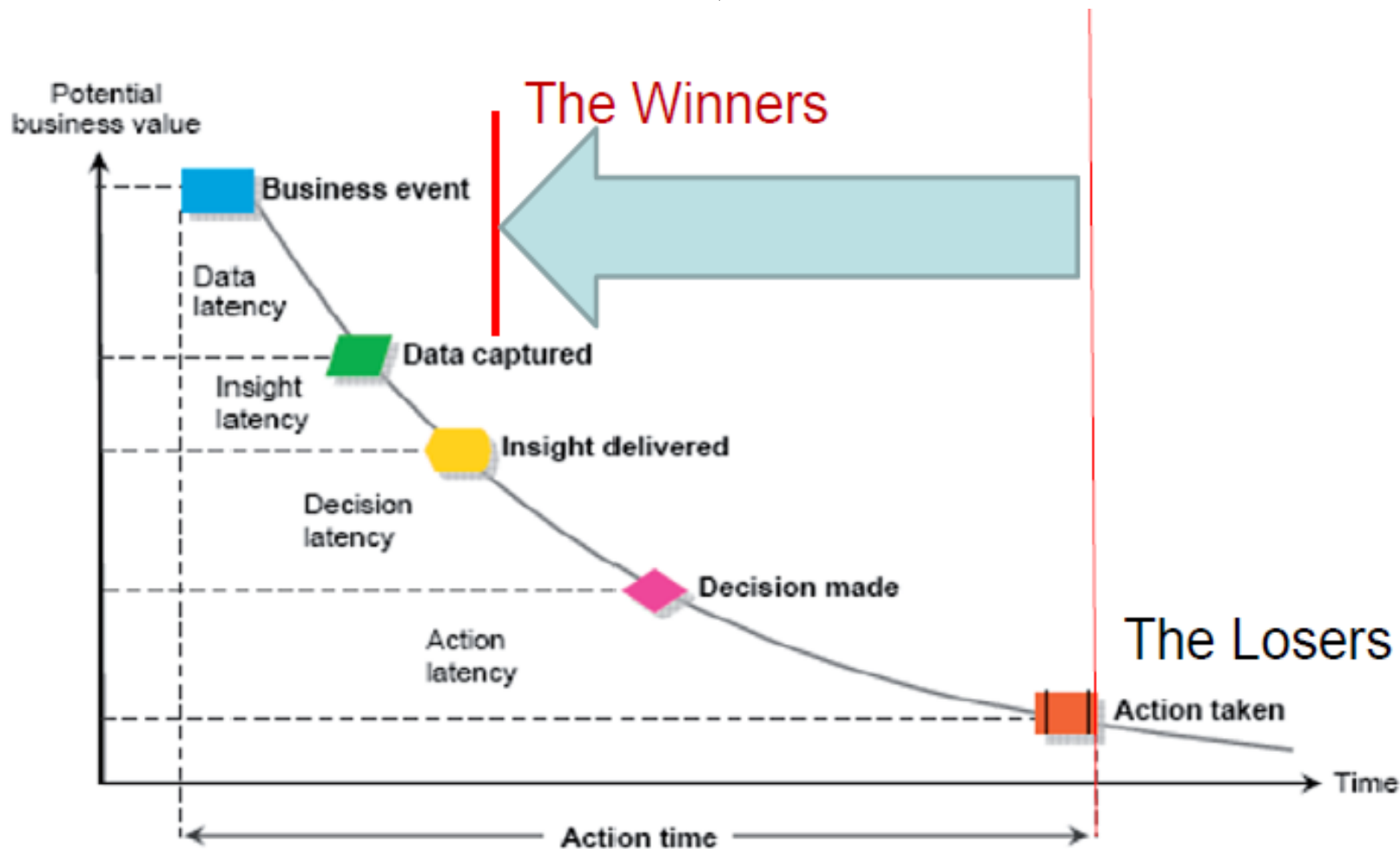


大數據的探採(mining)

可以找到意想不到的關聯性

- Story of Walmart in '90
 - Put beer and diaper together on the shelf
 - Will increase sales, especially on Friday
- Story of Target in 2012 (NY Times)
 - Angry father: why do you mail coupon for baby-related merchandize to my daughter, who is not 18 yet.
 - Grateful father: thank you for telling me my daughter is probably pregnant

大數據讓奧巴馬總統贏得選戰， 在各行各業都將如此

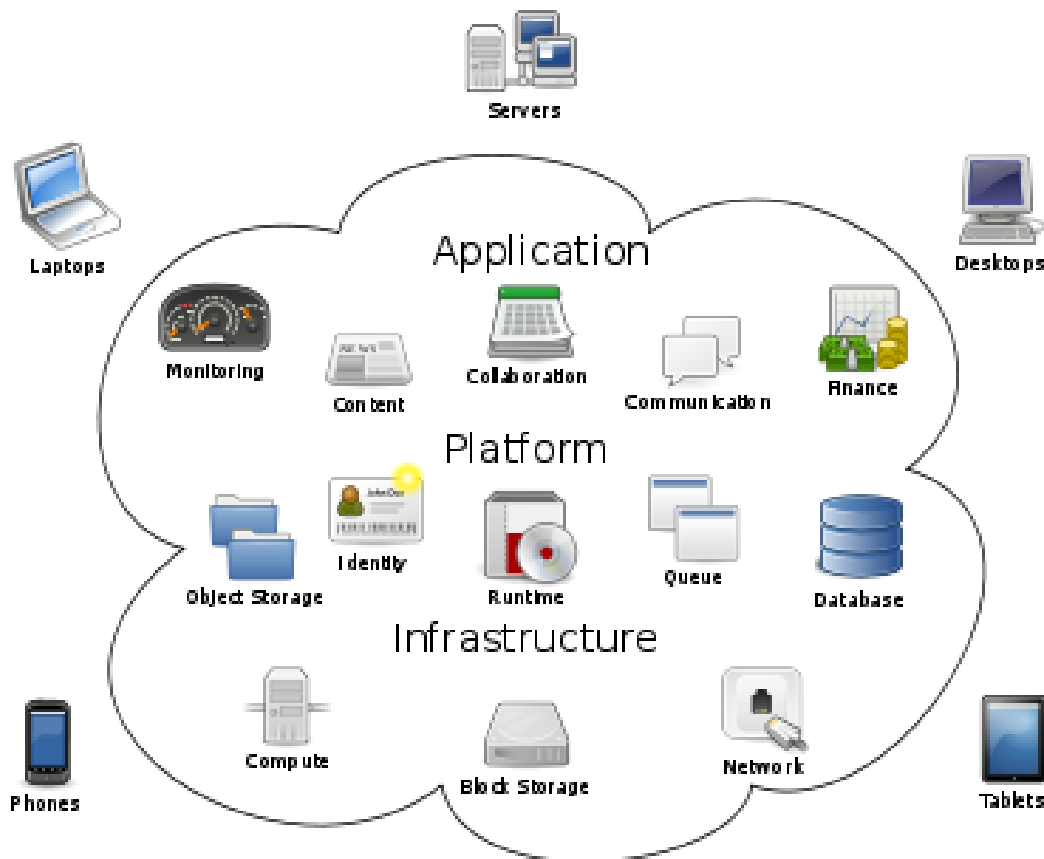


王可言(2013)全球物聯網技術與智慧產業發展趨勢

行動裝置的輸入不方便，很難做複雜的事，所以**需要與電腦相通**才會方便

- Mobile internet
 - Automation of knowledge work
 - The internet of things
 - Cloud technology (雲端技術)
 - Advanced robotics
 - Autonomous and near-autonomous vehicles
 - Next generation genomics
 - Energy store
 - 3D printing
 - Advanced materials
 - Advanced oil and gas exploration and recovery
 - Renewable energy
- by McKinsey Global Institute, May, 2013

雲端技術提供電腦與行動裝置間 跨平台的數據儲存與交換空間



只要將檔案儲存在 Dropbox 上，便可從電腦、手機、或平板電腦任意存取。在任何地方都能編輯文件、自動匯入相片、和分享影片。

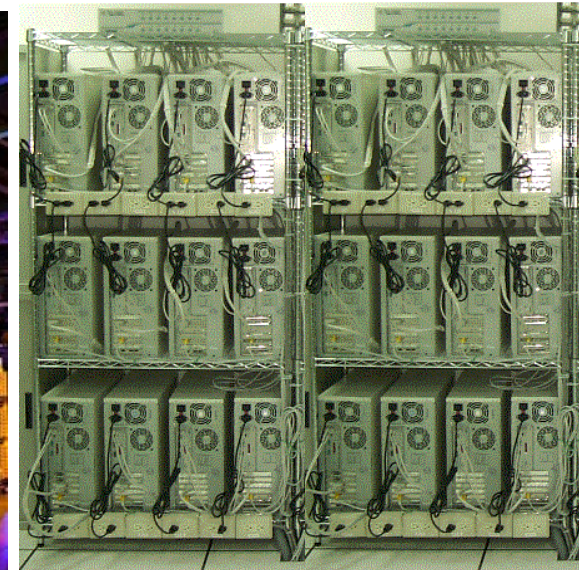
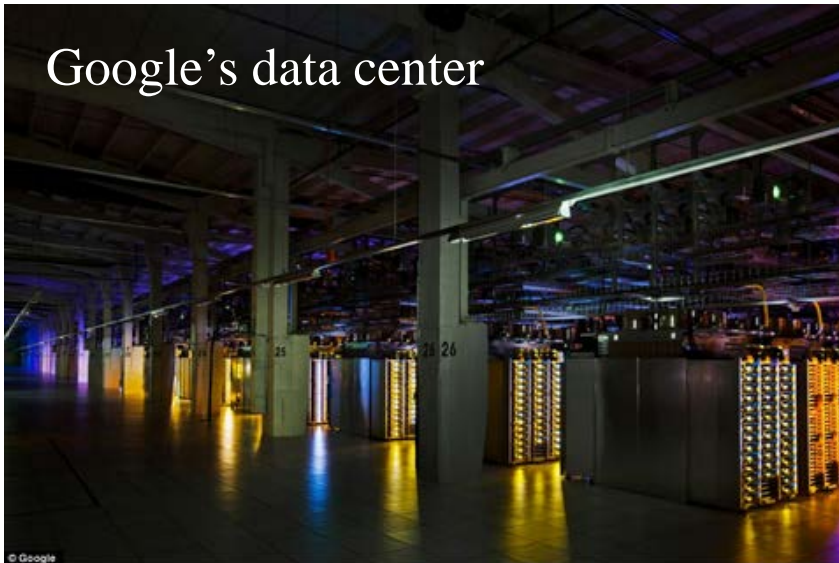
<https://www.dropbox.com/>

Cloud Computing
<http://earthnet.net/cloud.html>

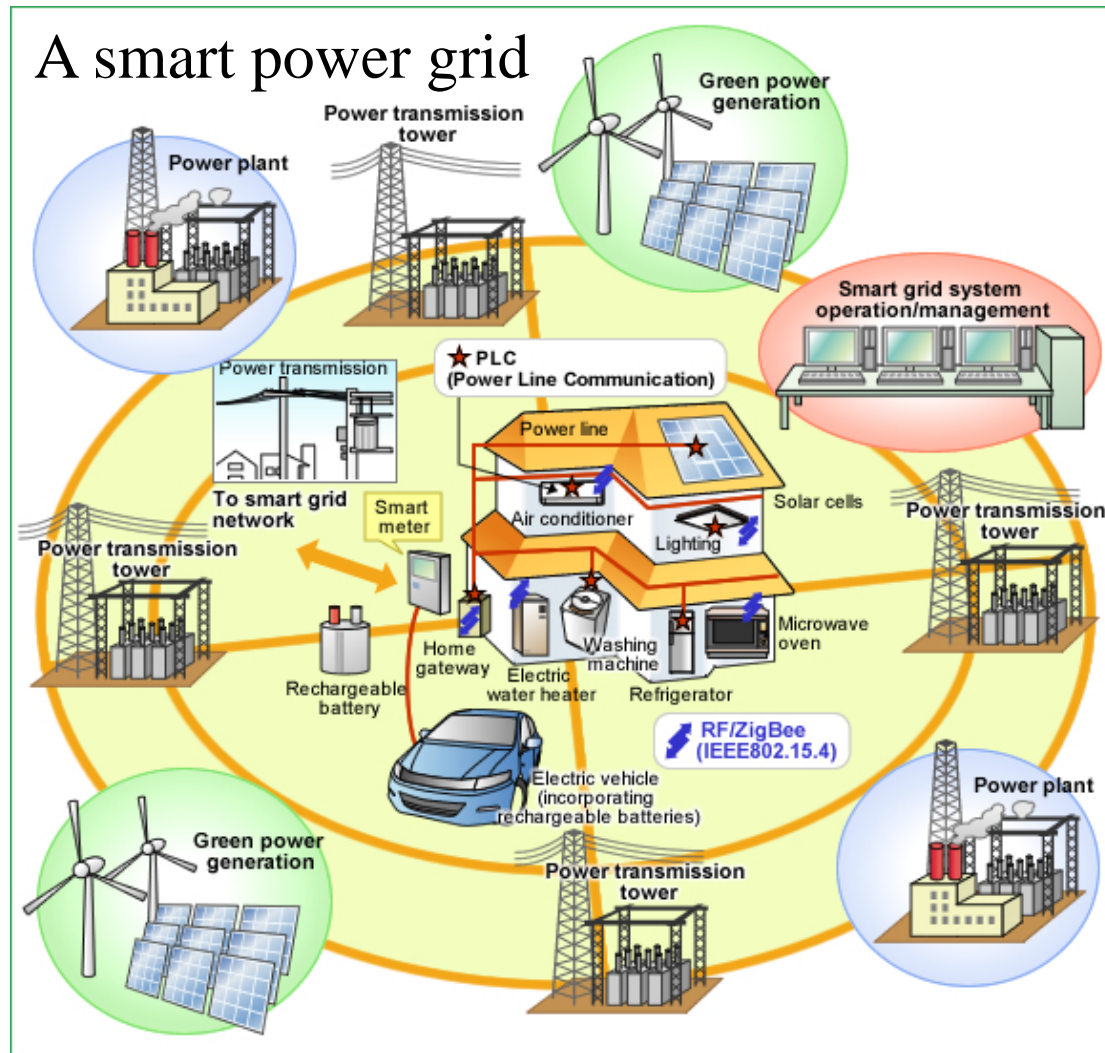
哪一種比較省錢？



mercedes-benz-s-class-production



當使用計算力像是用電，需要時打開，不用時關掉，有必要自己擁有發電廠嗎？



有字天書記載生命的故事，但是一種米 養百樣人，人人都有自己的故事

- Mobile internet
 - Automation of knowledge work
 - The internet of things
 - Cloud technology
 - Advanced robotics
 - Autonomous and near-autonomous vehicles
 - Next generation genomics (新世代定序)
 - Energy store
 - 3D printing
 - Advanced materials
 - Advanced oil and gas exploration and recovery
 - Renewable energy
- by McKinsey Global Institute, May, 2013

Angelina Jolie and genetic testing

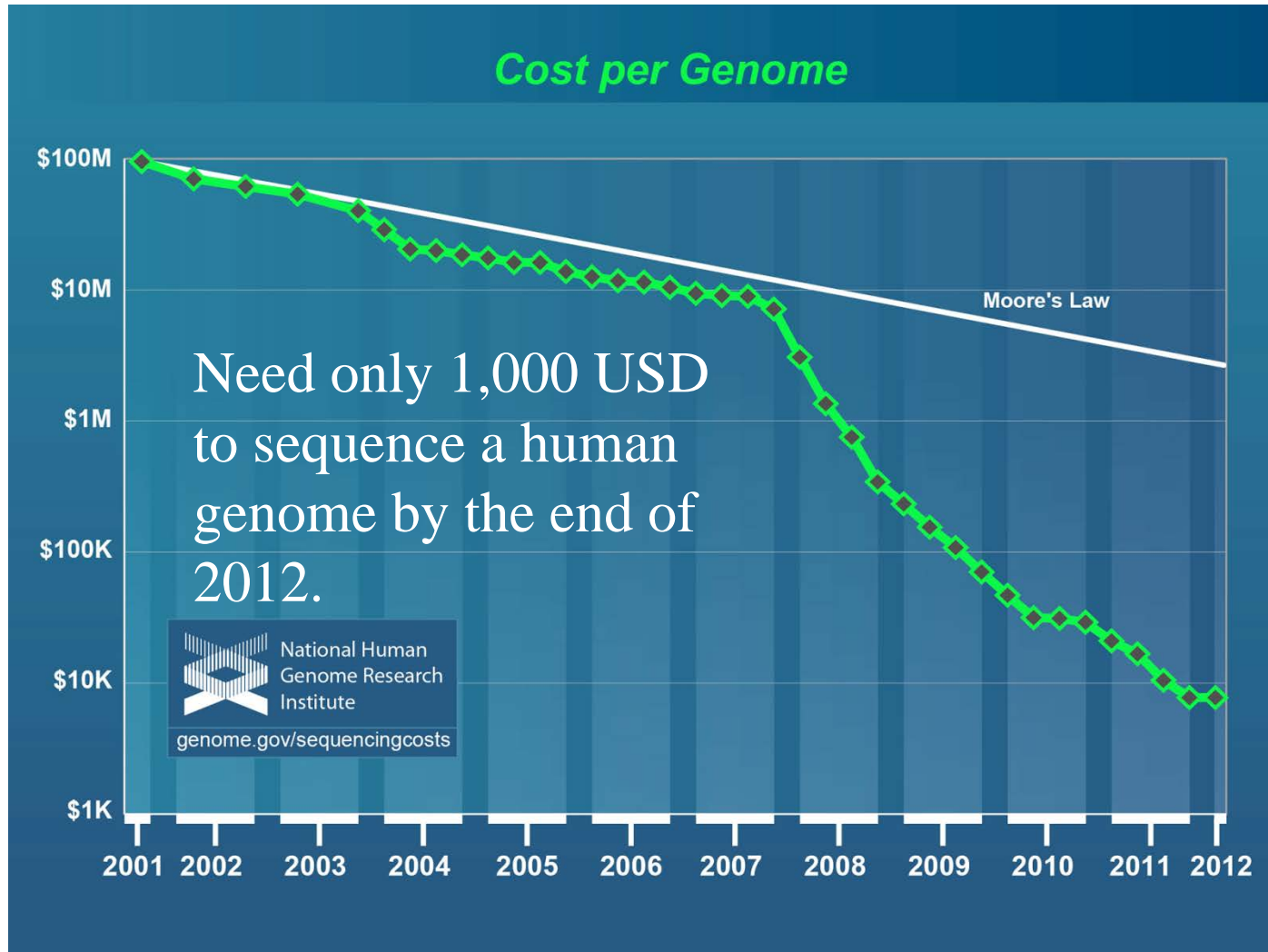
Table 1 Direct-to-consumer marketers of genetic information

Company	Technology	Cost
23andMe	Illumina HumanHap550+ BeadChip, oligonucleotide bead arrays in microwells on bundled optical fibers for detecting 550,000 SNPs plus 30,000 additional SNPs selected by 23andMe	\$399
deCODEme	Illumina Human1M BeadChip, for detecting >1 million SNPs	\$985 introductory price
Knome	Illumina genome sequencing platform	\$350,000
Life Code	Undisclosed SNP genotyping arrays	€1,200 to €2,400
Navigenics	Affymetrix Genome-Wide Human SNP Array 6.0, a photolithographically synthesized oligonucleotide chip for detecting 600,000 SNPs	\$2,500 plus \$250 per year "for continuous service"

Fox, J.L., Nat Biotechnol, 2008. **26**(10): p. 1105-8.

http://www.nytimes.com/2013/05/14/opinion/my-medical-choice.html?_r=0

定序價格急速滑落，讓我們有機會看到人的差異，也造成鉅量數據的累積



定序儀越來越小，最終將 進入醫院的檢驗中心

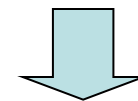


為什麼有些人吃藥會有嚴重的副作用？甚至造成藥害



Carbamazepine是抗癲癇與治療三叉神經痛用藥，有些人吃後，會引發史蒂文生氏強生症候羣(SJS)。

帶 HLA-B*1502 變異的病人，得到SJS的風險甚至是帶正常基因者的1千5百倍，嚴重者會造成死亡。

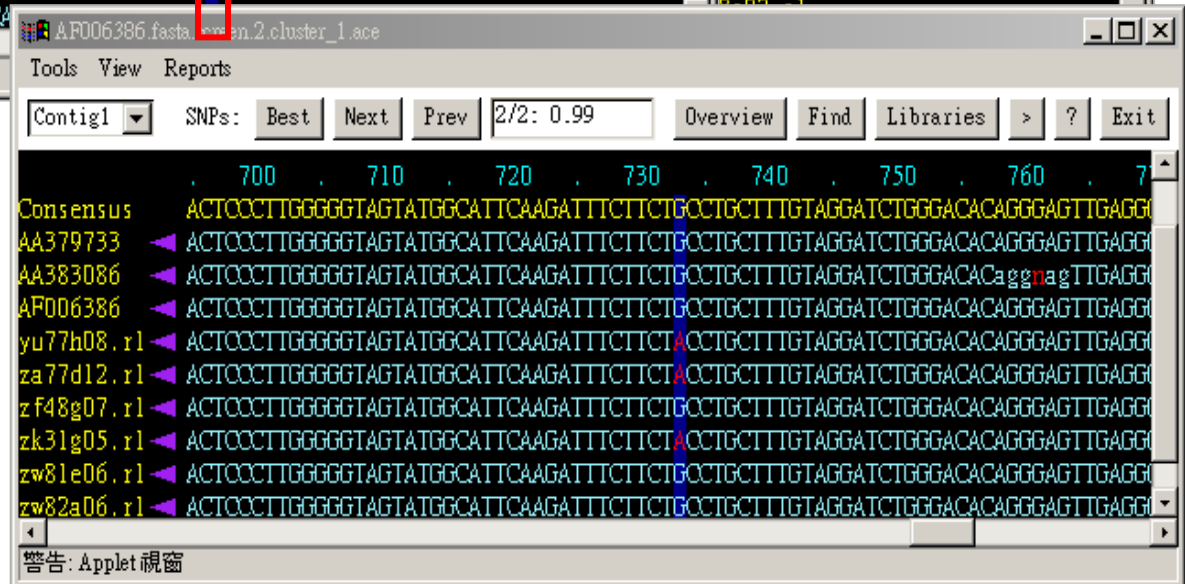
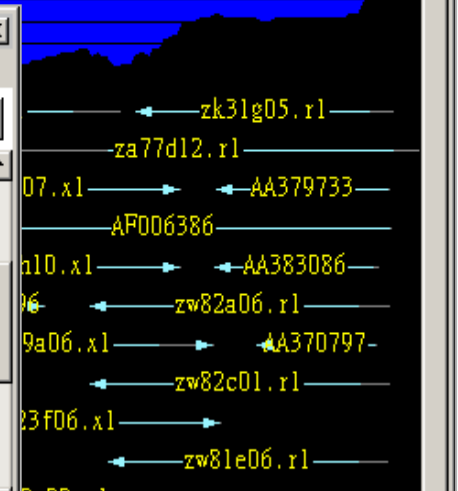
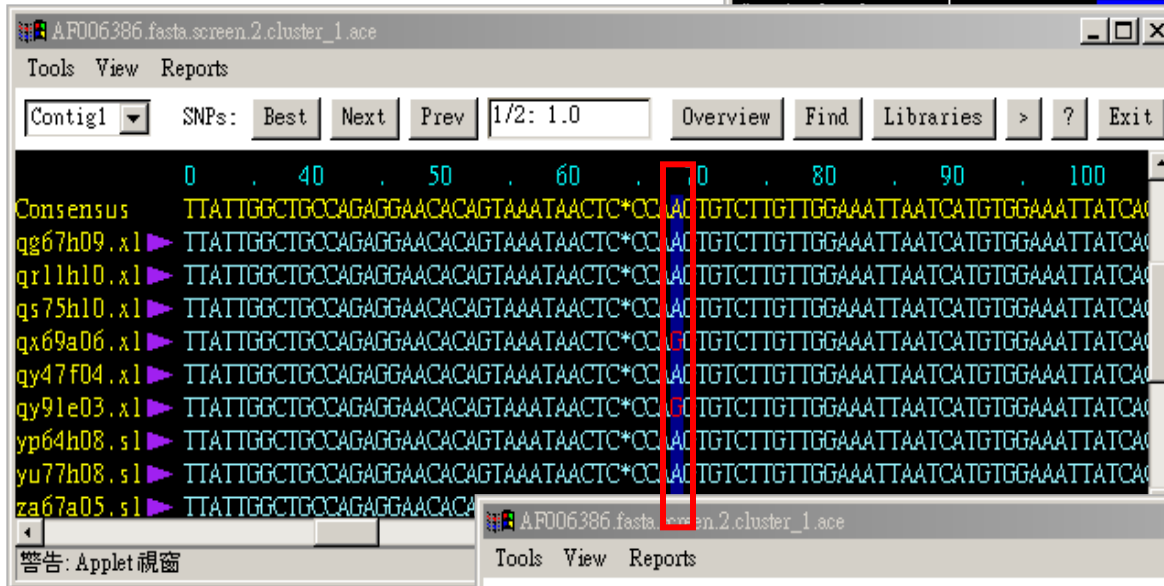
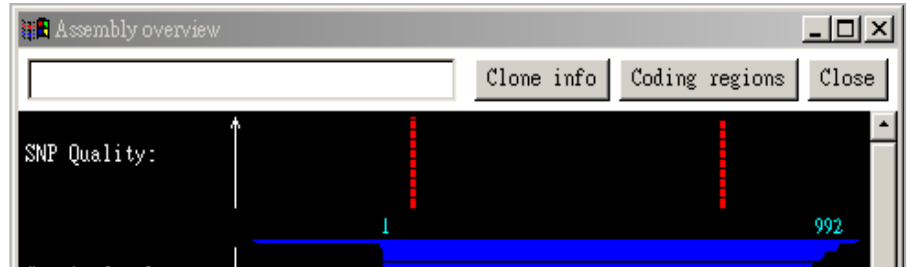


人的體質不同

Picture was taken from

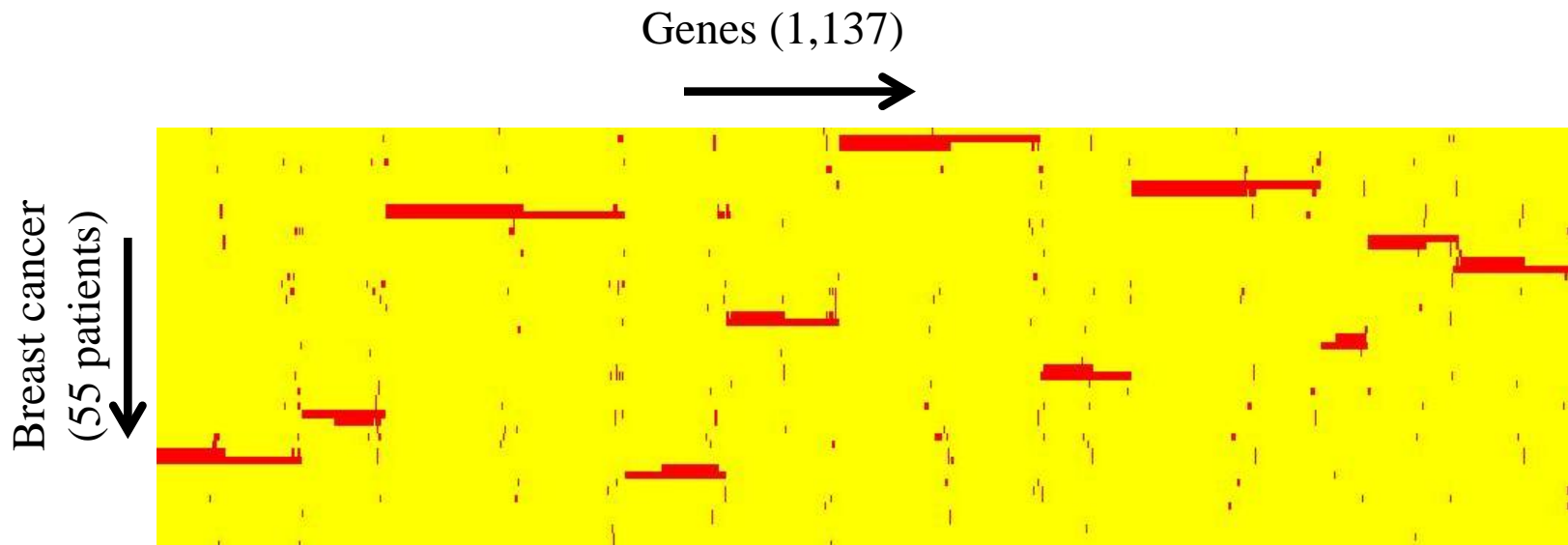
<http://www.health-pic.com/erythema-multiforme-stevens-johnson-syndrome/>

人與人的序列不盡相同，造成體質差異



許多變異是「中性」的，不會造成嚴重後果，只會使酵素活性變化，有時甚至完全沒有造成影響

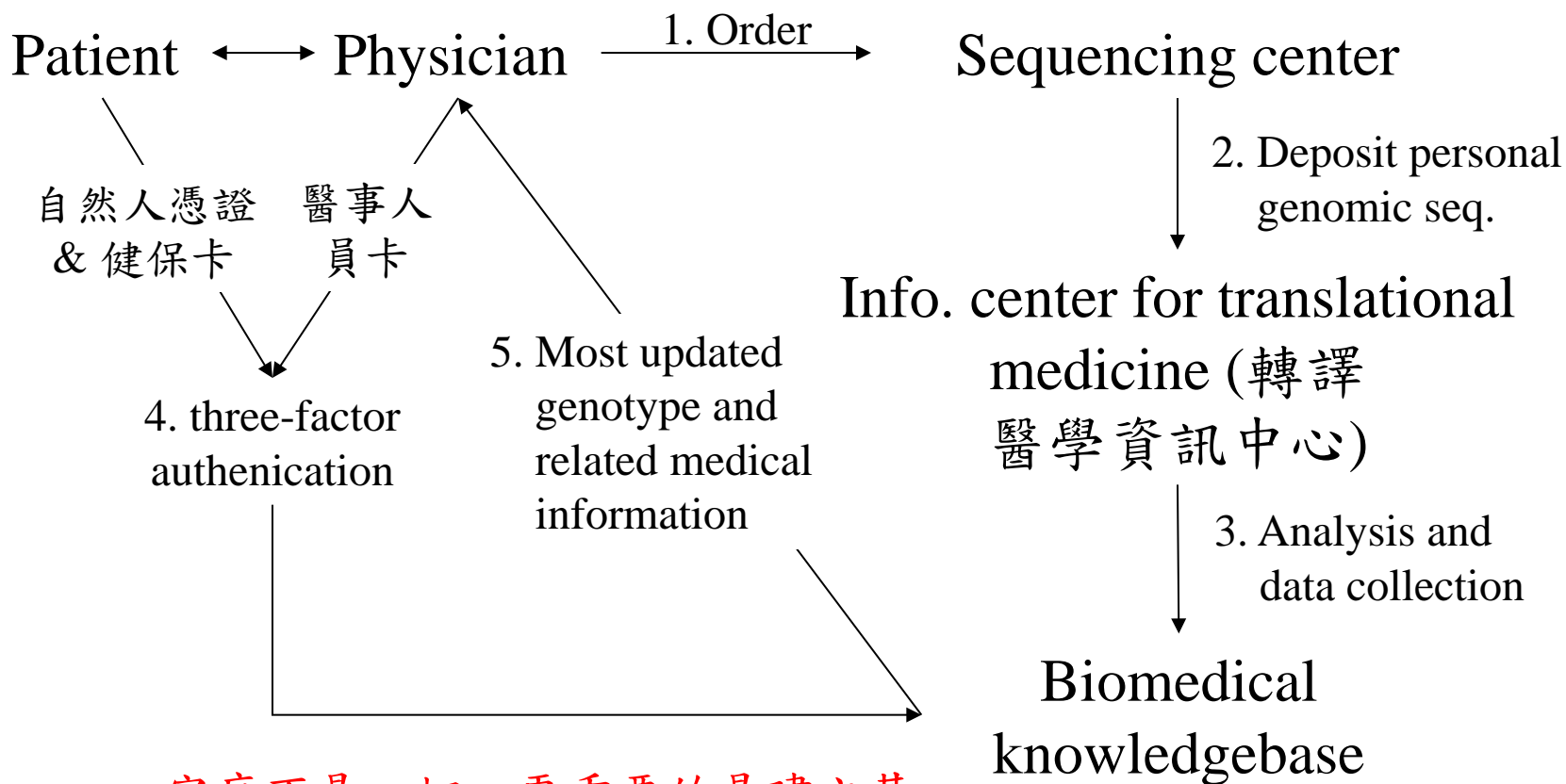
同樣是乳癌，每個人的腫瘤 卻有不同的突變



Plotted based on the data described in
Science, 314, 268-74 (2006)
Science 318, 1108-13 (2007)

個人化的醫療將不再是夢

(未來基因研究將成為醫學的一部份，而不再是另一學門)



定序不是一切，更重要的是建立基因與疾病，或基因與藥物間的關係

四種破壞式革新技術的衝擊

- Mobile internet (移動互聯網)
 - 開放資料與連結性 => 帶動「利用數據」的習慣
- The internet of things (物聯網)
 - 提供收集多樣資料的潛力 => 帶動「循證決策 (evidence-based decision making)」
- Cloud technology (雲端技術)
 - 什麼都要能自動化 => 帶動「流程改造」
- Next generation genomics (新世代定序)
 - 高解析度的測量，帶我們進入微觀的世界

越來越多的疾病被發現與遺傳有關

人類基因組計畫讓遺傳學研究的解析度大幅增加

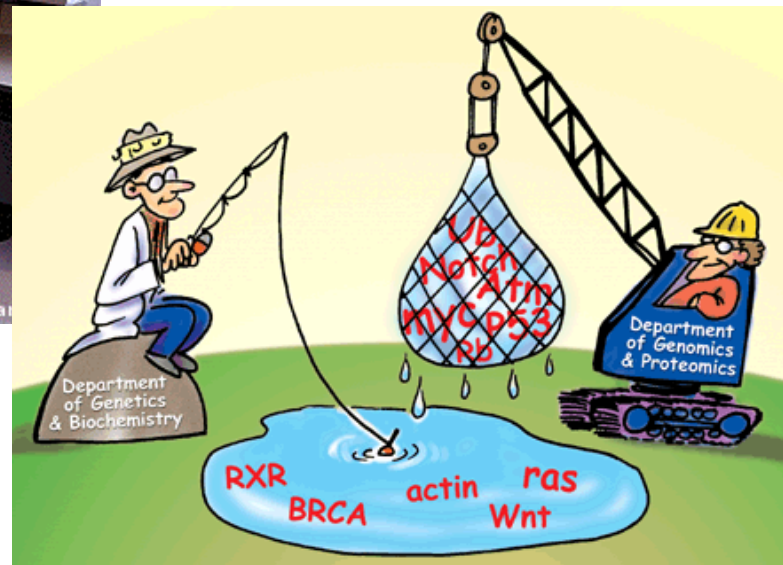


簡單的疾病：致病基因
複雜的疾病：危險因子

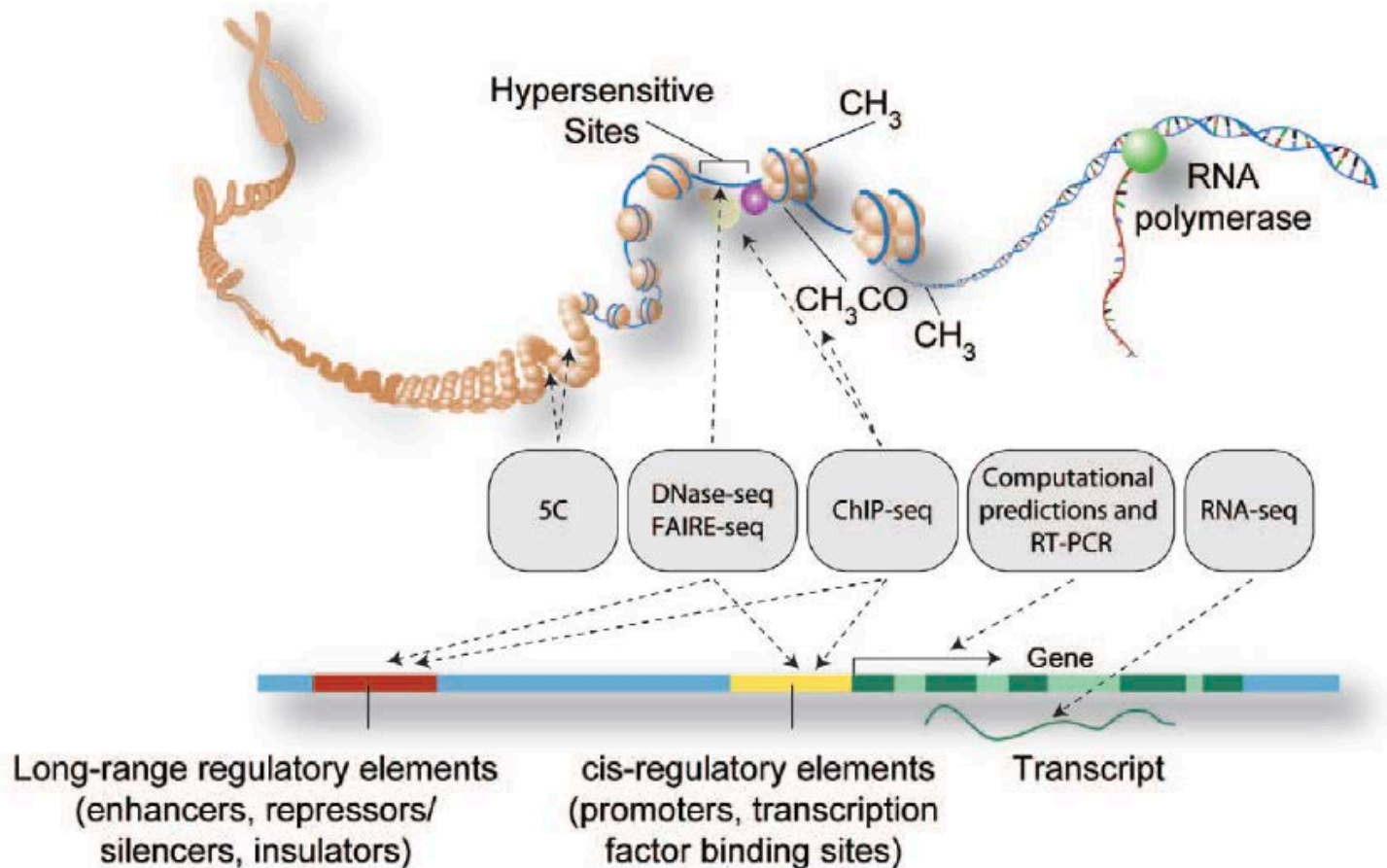


Left and middle figures: http://www.samogden.com/Human_Genome_Project.html

Right figure: <http://www.sciencemag.org/cgi/content/full/291/5507/1221/F1>

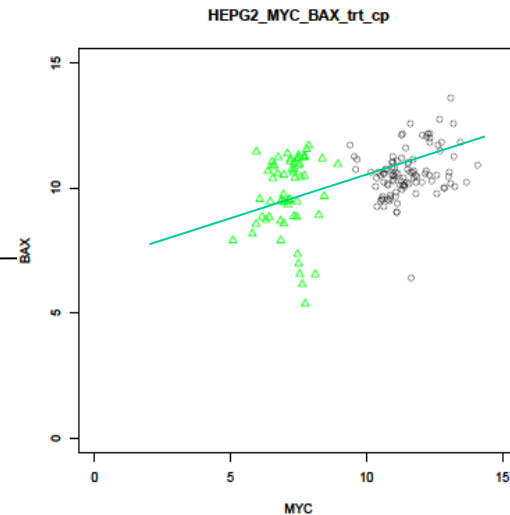
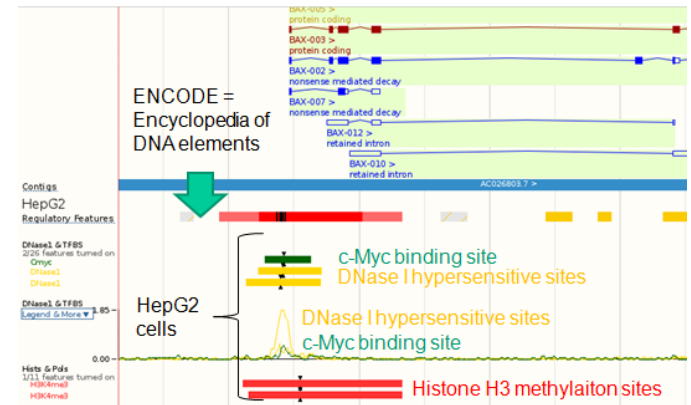


有系統地收集同一個體的各種相關資訊，
才能避免個體間差異造成的影響

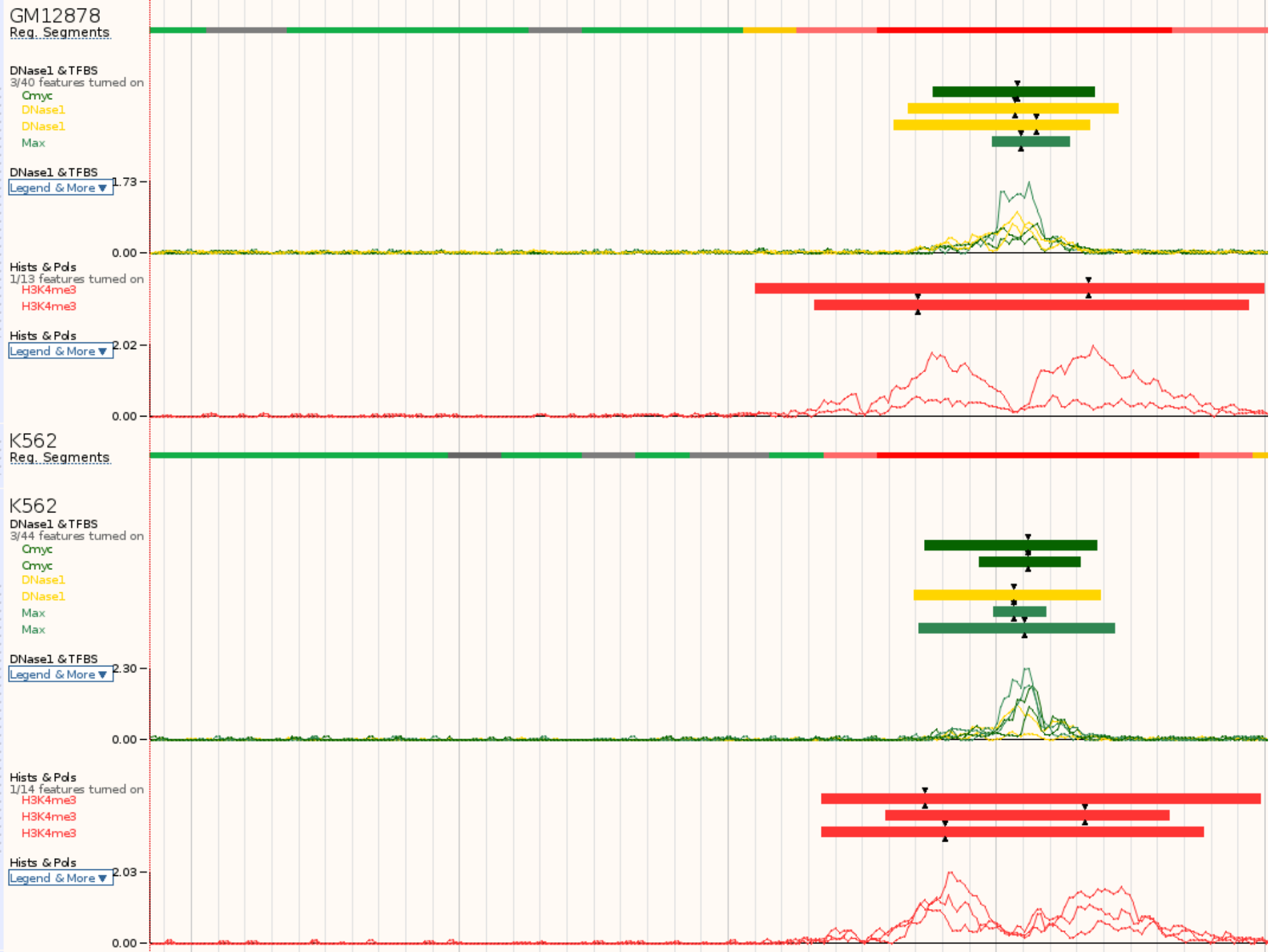


沒有一個團隊有那麼多錢， 可以收集所有的資訊

- The genomes
- Encyclopedia of DNA elements (ENCODE)
- The cancer genome atlas (TCGA)
- Library of Integrated Network-based cellular signatures (LINCS)
- Stem Cell database (StemDB)
- ..., *etc.*



-
- National health insurance research database
 - Taiwan biobank



臺灣研究生物醫學的優勢： 全世界唯一的全民健保資料庫



全民健康保險研究資料庫
National Health Insurance Research Database



NEW 97年資料發行 **ENGLISH**

非學術界研究類

學術界研究類

學術界研究類

政府立案之國內公私立大專院校、政府所屬業務相關單位、或非營利研究機構（含教學醫院）之講師、技正（或相當職等）、助研究員、專科醫師等（含）以上或其他經「全民健康保險研究資料庫指導委員會」（以下簡稱指委會）審核通過者。

National Health Insurance Research Database

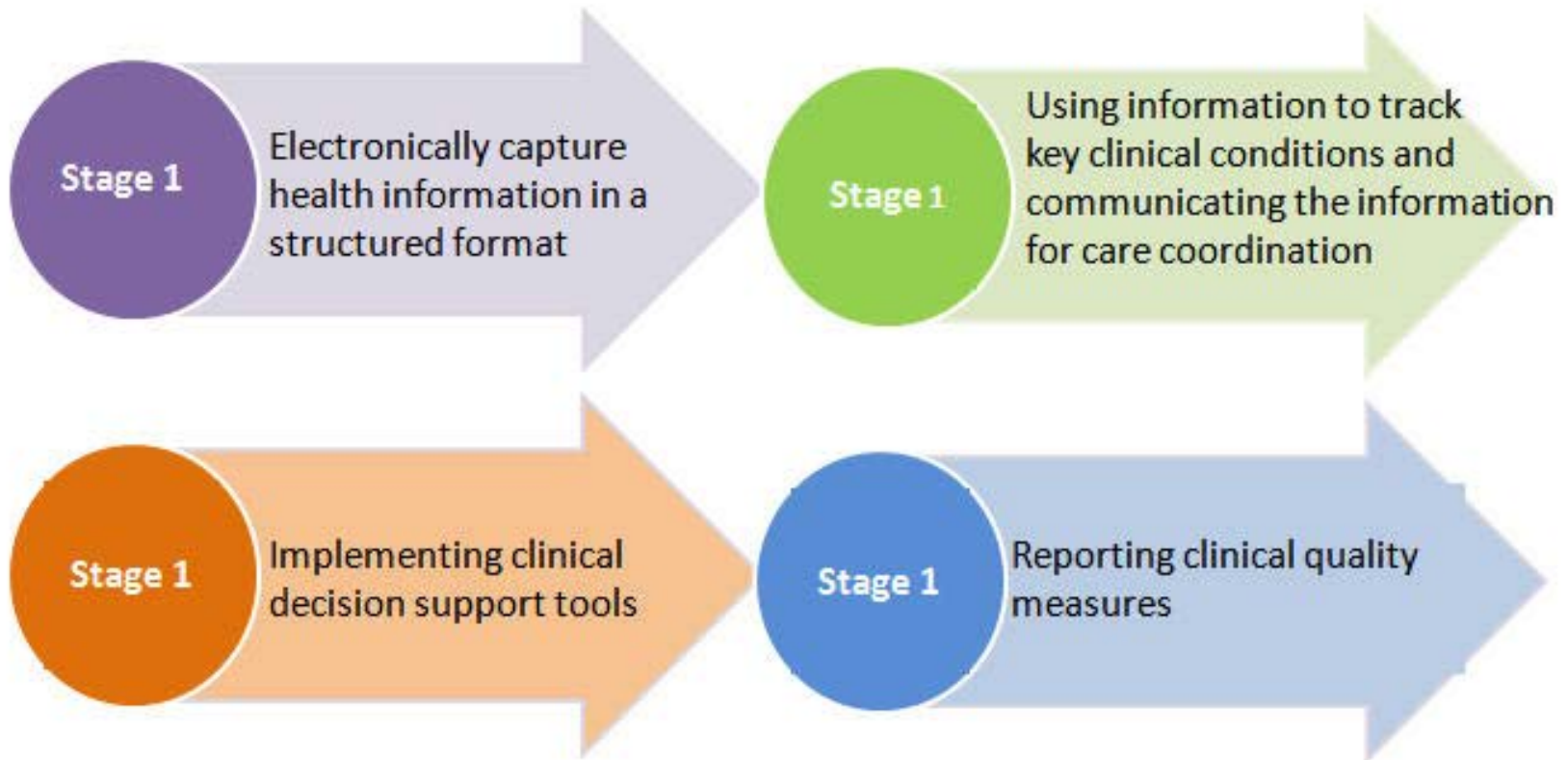
最佳瀏覽效果800x600

Copyright 2003 National Health Research Institutes. All rights reserved.

<http://w3.nhri.org.tw/nhird/>

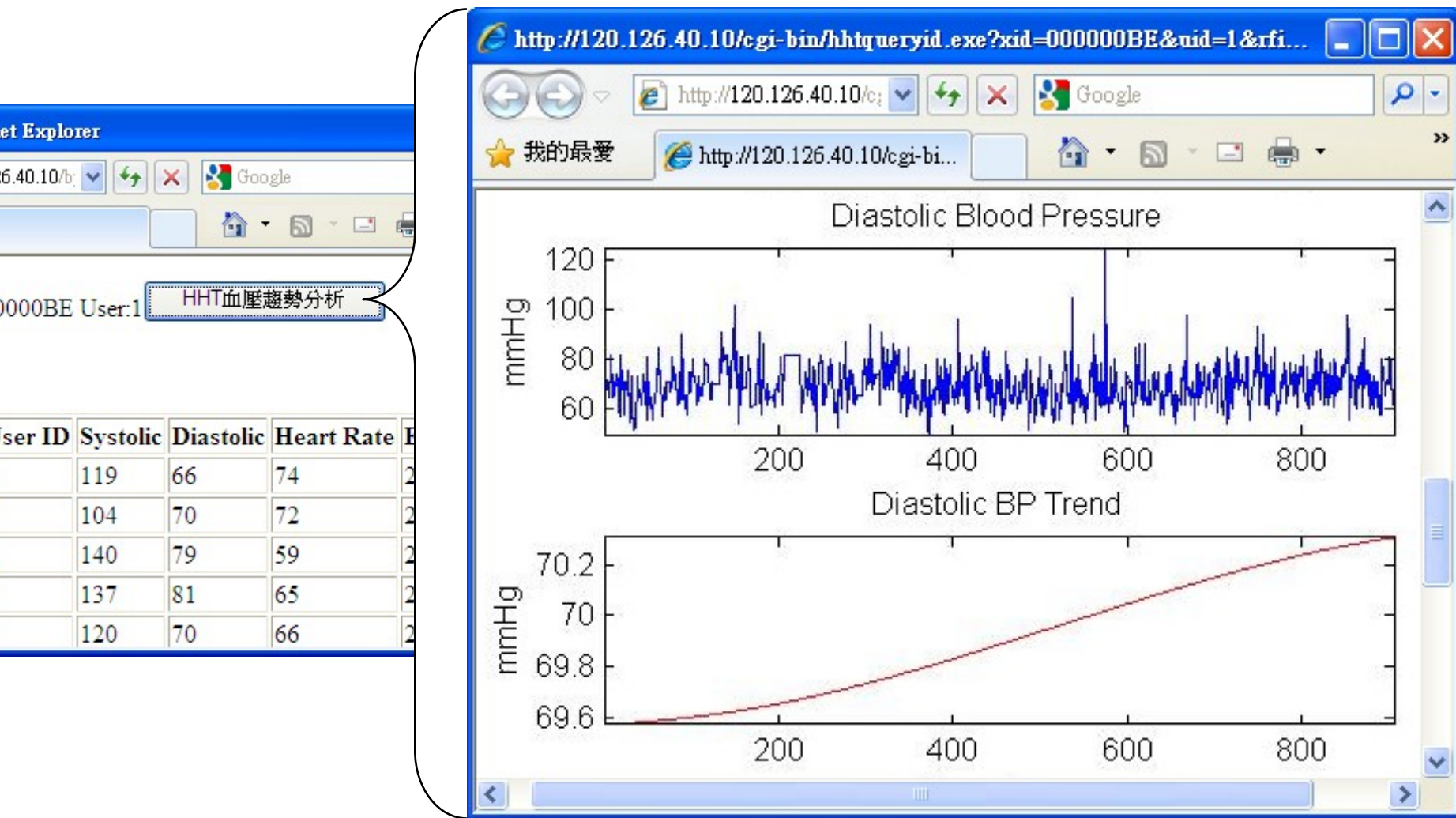
電子病歷救經濟？

Meaningful Use Criteria – Stage 1



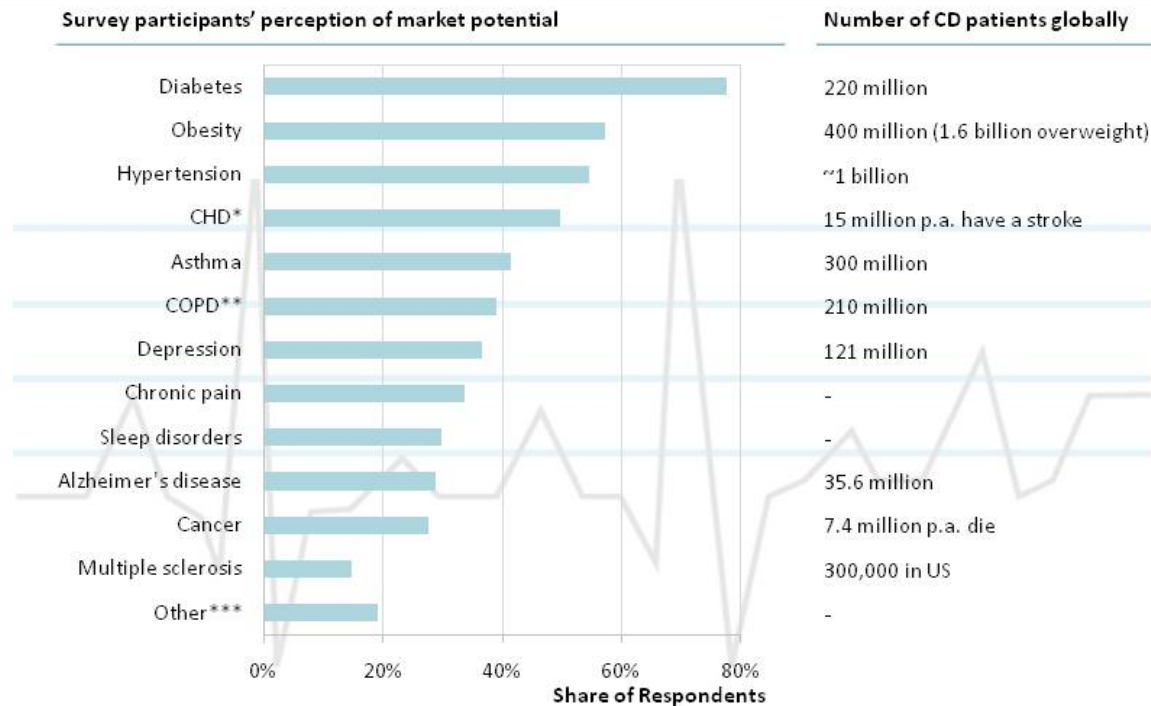
<http://www.cntit.com/healthcare-solutions-practice-management-emr-government-emr-incentives.htm>

個人的健康數據，將隨著物聯網技術的進步，而越來越有**應用價值**



移動健康 (mobile health, mHealth) 將成為生活的一部份

Therapy Fields That Offer the Best Market Potential for mHealth Solutions

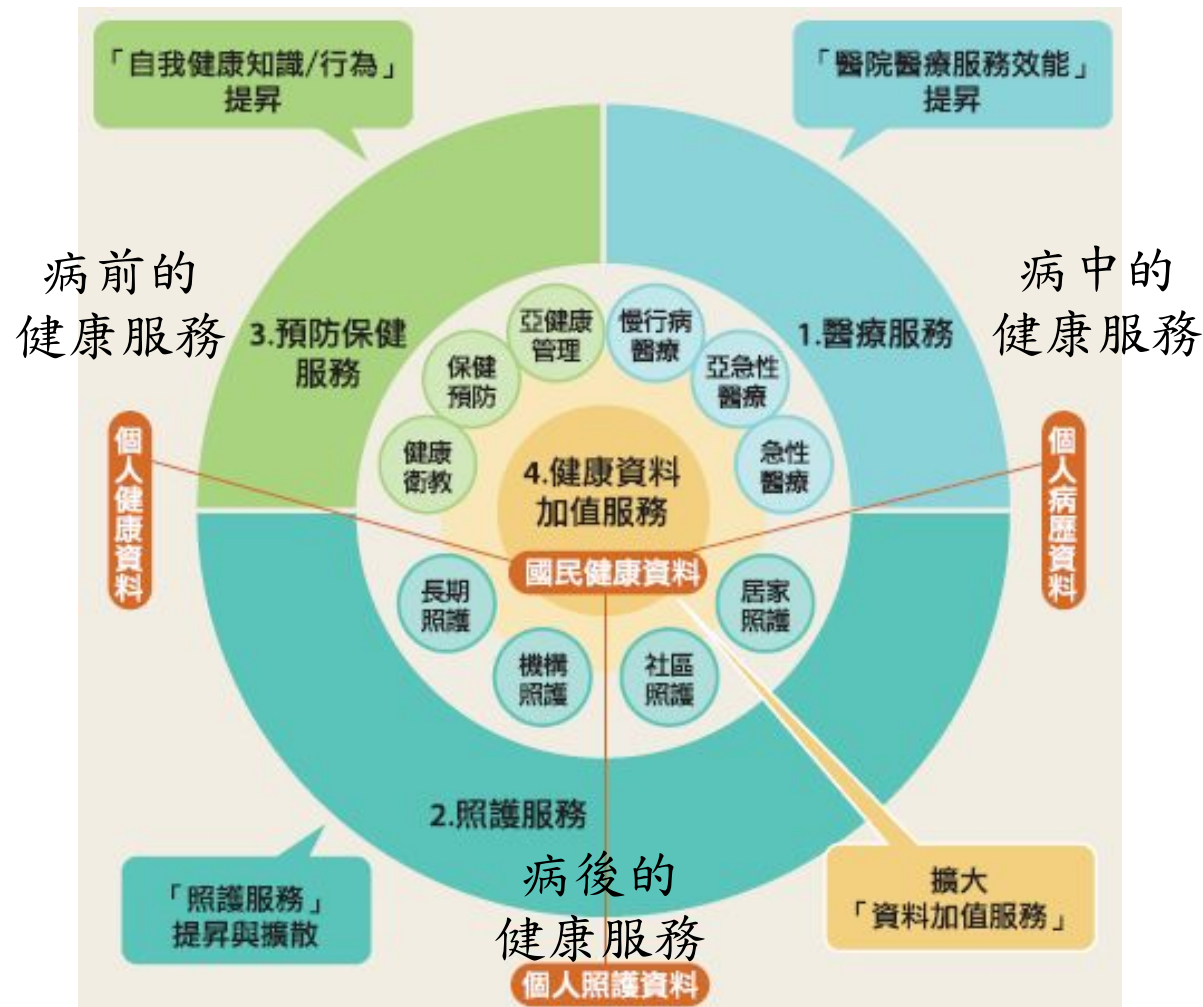


* CHD refersto coronary heart disease. ** COPD refersto chronic obstructive pulmonary disease
 *** Other chronic diseases such as HIV, Aids, epilepsy, infectious diseases, other less serious illnesses
 Source: research2guidance Global mHealth developer survey , n= 231; OECD, WHO data.

research2guidance
 the mobile research specialists

<http://www.research2guidance.com/global-survey-results-220-million-diabetes-patients-are-the-main-target-for-mhealth-app-developers-in-the-next-5-years/>

台灣的健康雲計畫包括 保健、醫療、與照護三朵小雲



<http://www.ithome.com.tw/itadm/article.php?c=77236>, 2012-11-14

你放心將健康數據放在雲端嗎？

道高一尺，魔高一丈

- Hundreds of intimate pictures of female celebrities, including Oscar-winning actor Jennifer Lawrence, were leaked overnight after being stolen from their private collections and posted on infamous web forum 4chan.
-
- Final Destination actress Mary Elizabeth Winstead - another of the hacking victims - claimed that she had deleted her intimate shots years earlier, suggesting that fully removing iCloud images may be more complicated than previously understood, or that the hacker had been compiling the images for some time.

<http://www.newsweek.com/hundreds-intimate-celebrity-pictures-leaked-online-following-suspected-icloud-267851>



天下沒有不可破解的技術

病人需要有隱私

圖片取自 <http://www.comicmana.com/html/managallery/ysl/ysl0005.htm>

立委林進興等12位醫師公佈胡志強病歷，台中市政府醫師懲戒委員會，做出廢照一年，或是接受醫學倫理教育的處分，但有八成網友認為這樣的處分太輕了！

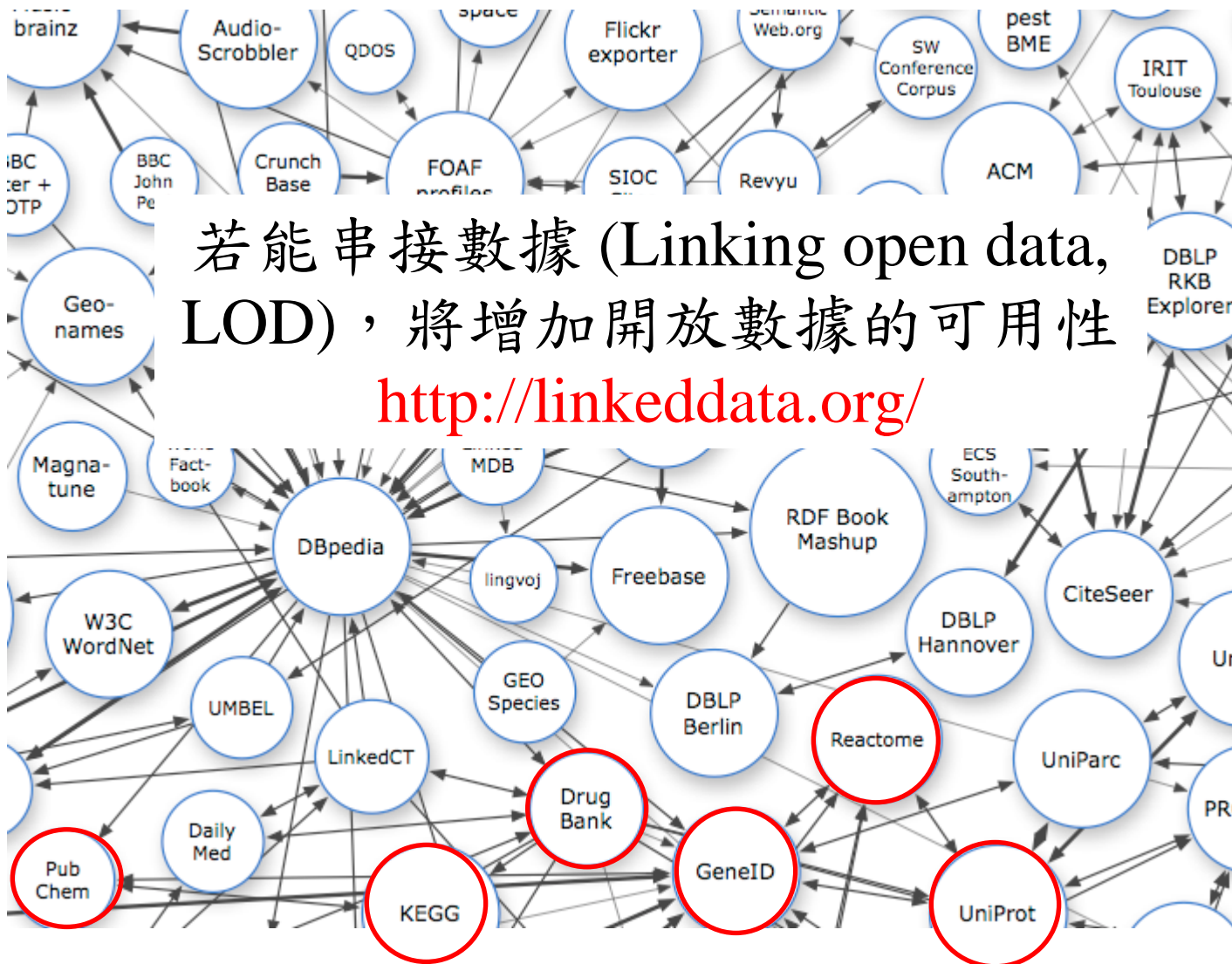
(http://www.tvbs.com.tw/news/news_list.asp?no=alisa20051209202635)



圖片取自 http://www.tvbs.com.tw/FILE_DB/newsphoto/alisa/200512/alisa-20051209214847.JPG

必須培養從業人員
的職業道德
與全民的道德觀

生醫領域的開放數據多數可以串接

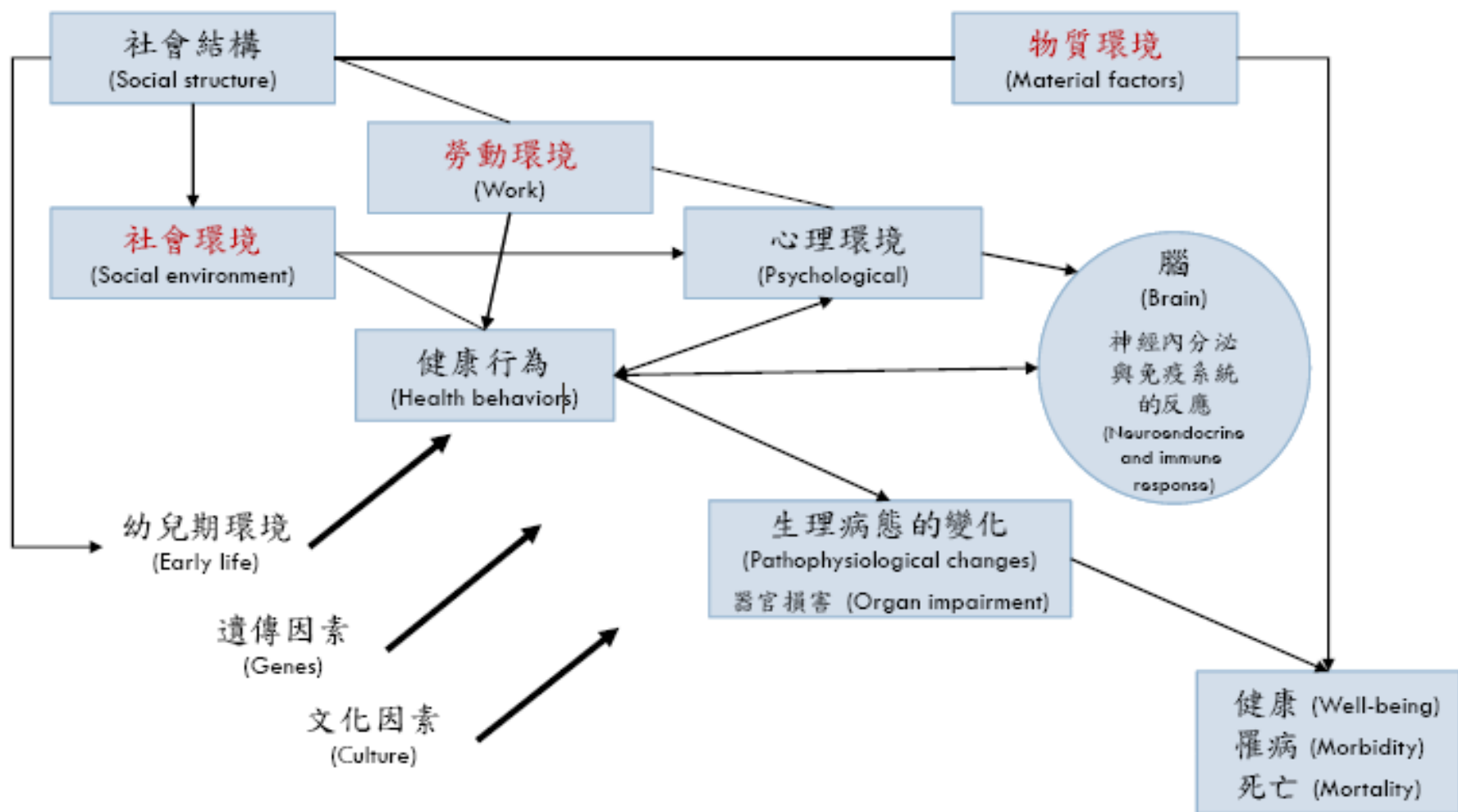


串接資料可以回答許多問題

- 例如研究冠狀動脈繞道手術的醫療品質研究時，發現醫師手術死亡率與其手術量呈負相關
 - 方法(一): 取得「可整合的」健保資料與死因檔後，根據需要自行整合
 - 方法(二): 由健保資料庫中，取得醫師與病人間的關係，只要求部份(指定)病人的死因
 - 是否有第三種方法，例如**虛擬整合**等，則有待研究

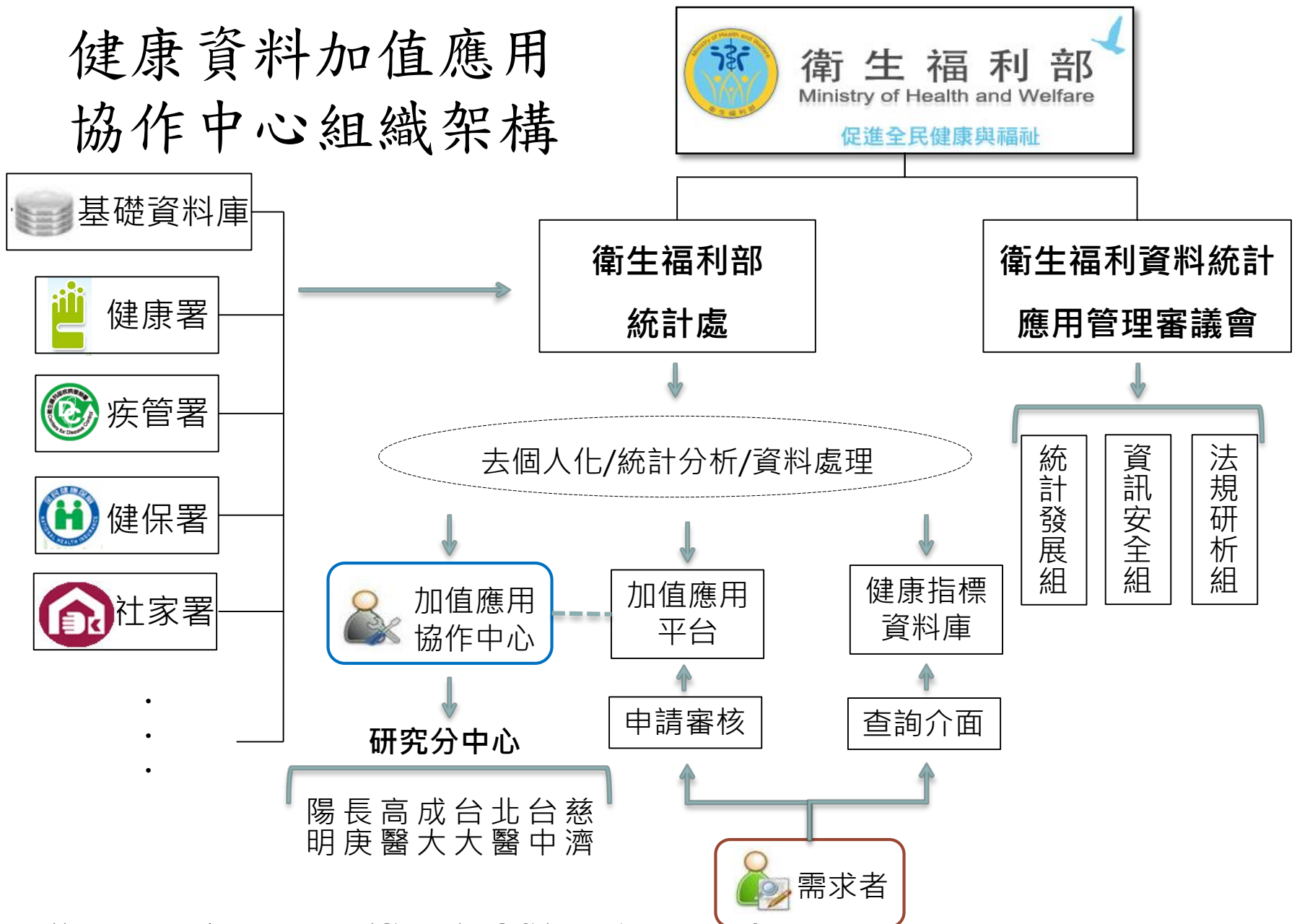
ELSI = ethical, legal, and social issues

親人有同樣的基因， 卻不見得人人得同樣的病



健康資料加值應用實例介紹,衛生署統計室,100.11.11

健康資料增值應用 協作中心組織架構



台灣人體生物資料庫透過Taiwan view網站，已開始釋出基因型相關資訊

English | 中文

Taiwan Biobank TWB Chip How to Use Search Registration Login

臺灣人體生物資料庫成立目的是結合生活習慣、環境因子與生物標誌等資訊，建立屬於臺灣本土的人體生物資料庫，為生物醫學研究蒐集龐大的生物檢體與健康資訊，提供國內各領域的研究學者申請使用。

目前許多基因醫學的研究採用全基因相關性 (Genome-wide association, GWA) 研究策略。相對於候選基因之策略，採用全基因體掃描 (Whole Genome Screen) 更能找出新穎基因對疾病的影響，也助於全面性評估疾病的遺傳貢獻。

本資料庫依研究趨勢、未來適切性、應用性、研究經費與成本考量，目前採用全基因體掃描檢測遺傳標記，基因鑑定資料之釋出將去除身份識別資訊，以確保達到保護參與者隱私權益，提供各領域的研究學者申請使用，進行描述性分析，作為基因醫學研究的基礎，以建立研究之初步假說。

News

- December, 2012 : TaiwanView database 2.0 now available.

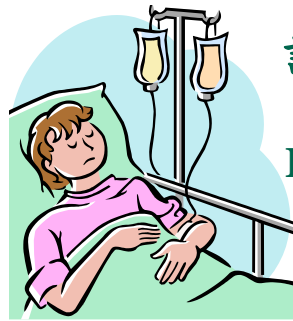
Database Information

Version: 2.0
Subject: 1802
Source: healthy controls
Update: 2012/12/17

以證據為基礎的醫學 (evidence-based medicine)

任何的處置都要依據證據，而臨床
試驗的結果比研究的推論更**直接**

由美國國家衛生院導入臨床資訊管理系統 (Clinical Informatics & Management System)



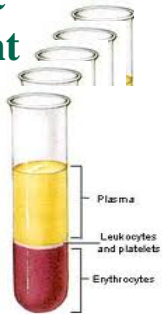
Patient

試驗紀錄需CSIS
(Clinical Study
Information System)



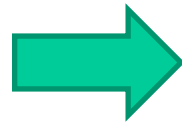
Clinical information

檢體管理需
STAMS
(Specimen
Tracking &
Management
System)



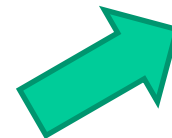
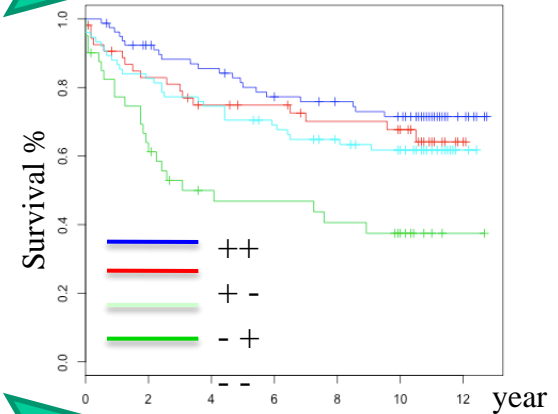
Specimen

計畫審核需PTMS
(Protocol Tracking
and Management
System) Informed consent



High throughput data

資料整合需
GUID (Global
Unique Identifier)



Analysis

<http://www.cims.tw/>

使用別人收集的數據，最怕有
「垃圾進，垃圾出」的情形

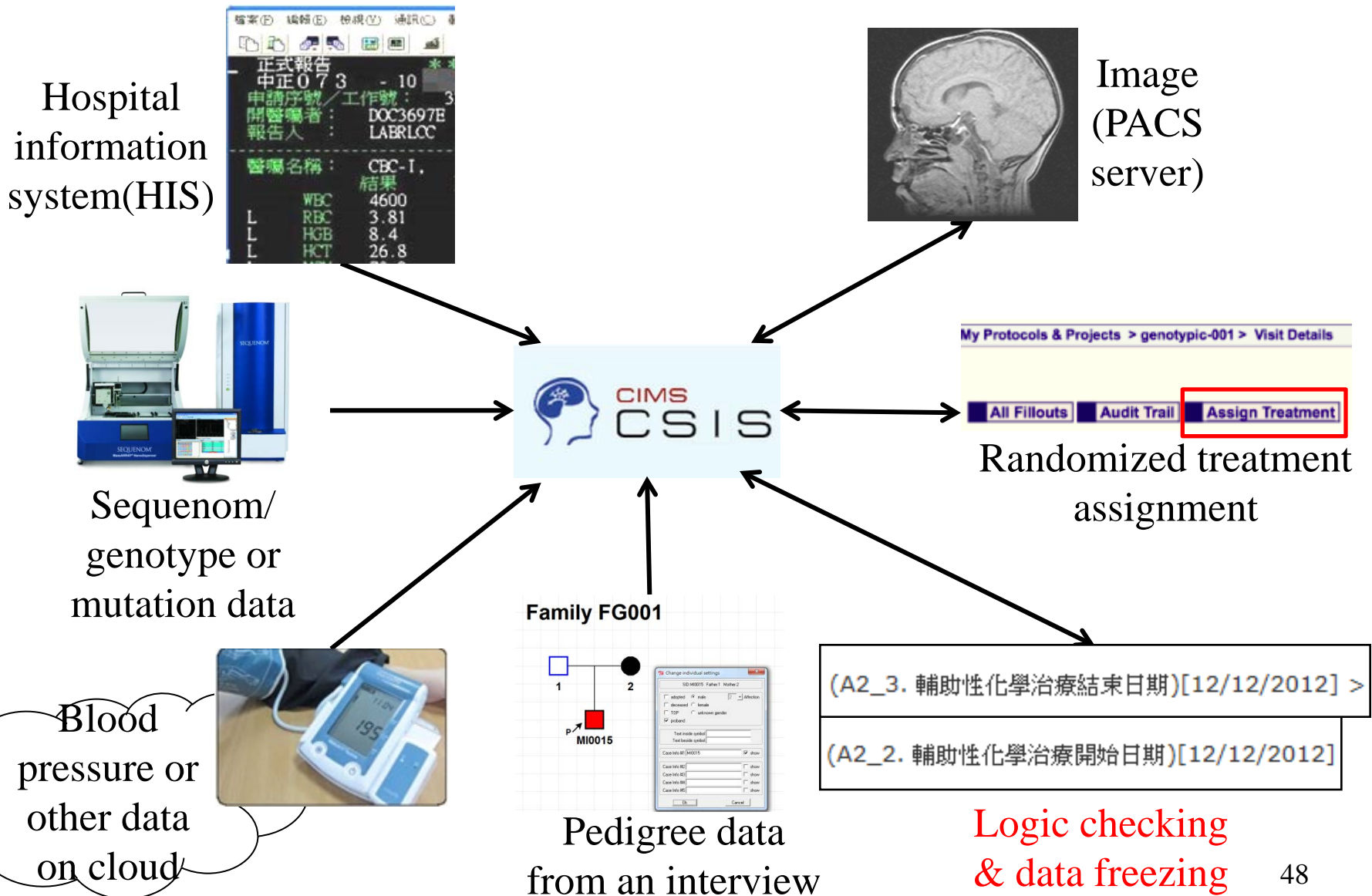
若要共享數據，必須有標準作業
程序，與品質管控機制

台灣已有超過80%的疾病臨床試驗聯盟使用CSIS

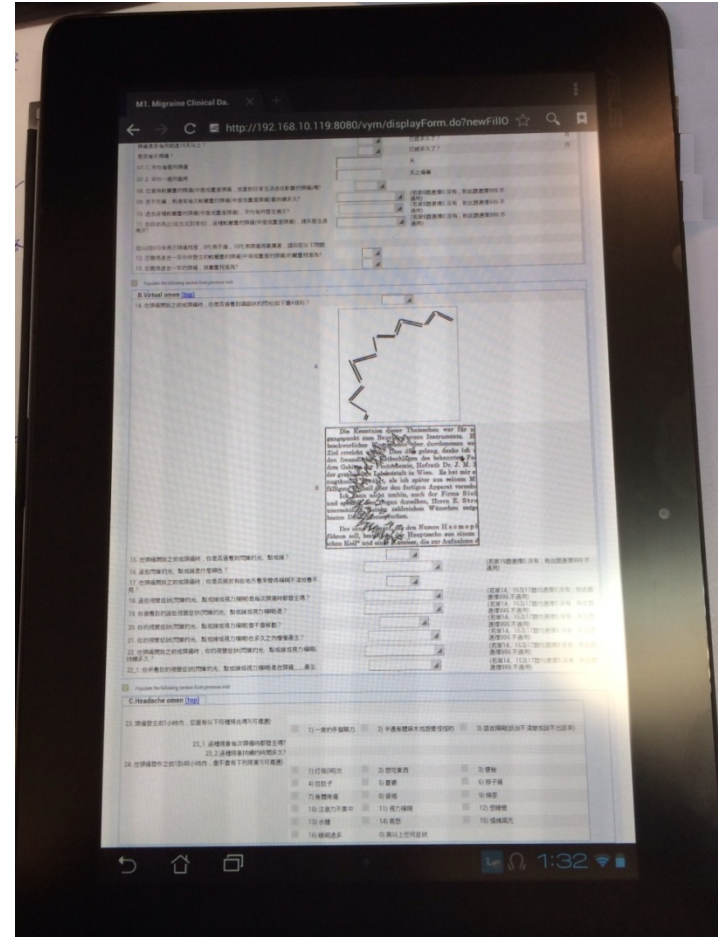
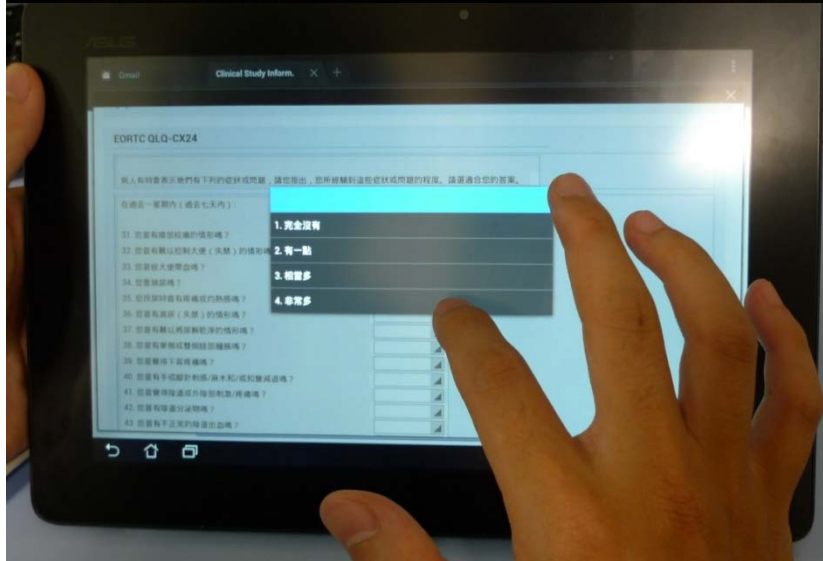
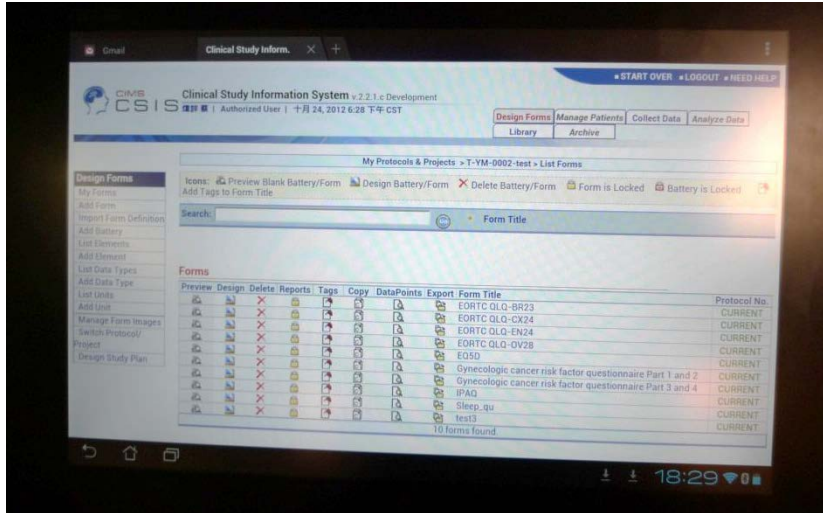
<http://tc2.ntu.edu.tw/en>

 <p>肺癌臨床試驗聯盟 Lung Cancer Clinical Trial Consortium ✓</p>	 <p>臺灣肝炎及肝癌臨床試驗聯盟 Taiwan LiverNet Consortium ✓</p>
 <p>癌症早期臨床試驗聯盟 Oncology Phase I Consortium ✓</p>	 <p>臺灣胃腸疾病及幽門桿菌臨床試驗聯盟 Taiwan Gastrointestinal Disease and Helicobacter Consortium ✓</p>
 <p>婦科癌症研究聯盟 Gynecologic Oncology Group ✓</p>	 <p>乳癌臨床試驗合作聯盟 Breast Cancer Consortium (✓)</p>
 <p>血脂和動脈粥樣硬化聯盟 Consortium of Lipid and Atherosclerosis ✓</p>	 <p>高血壓相關疾病試驗聯盟 Consortium of Hypertension associated Cardiac Disease ✓</p>
 <p>慢性阻塞性肺病聯盟 COPD consortium ✓</p>	 <p>法布瑞氏症試驗聯盟 Fabry Disease Consortium ✓</p>
 <p>小兒感染症聯盟 Pediatric Infectious Diseases Alliance</p>	 <p>精神疾病臨床試驗聯盟 Consortium for Mental Disorders ✓ √47</p>

臺灣版CSIS有多種確保數據品質的機制



在可以無線上網處， 即可使用行動裝置填寫問卷



檢體是連接臨床研究與基礎研究的橋樑

CIMS STAMS Specimen Tracking and Management System v1.0 Development
 Cimsuser Ptms | Lab Admin | Jun 10, 2009 1:23 PM EDT

Home Add Box Type Add Sample Types Config Sample Types Mapping

Main Menu Search & Export Reports Storage

Bldg. 10 Repository
 freezer1
 rack 1
 shelf 1
 box 1
 shelf 2
 shelf 3
 shelf 4
 rack 2
 freezer2
 freezer3

box 1/shelf 1/rack 1/freezer1

Box1中第一個檢體的資訊

[Edit Sample](#)
[Sample History](#)

Sample Information:

Position X 4
 Position Y 3
 Sample Created: Fri May 08 14:21:34 EDT 2009
 Sample Barcode 888888888888
 Date Collected: Fri May 08 14:21:34 EDT 2009
 Date Frozen: Fri May 08 14:21:34 EDT 2009
 Collect Units 0
 Volume 8899
 Sample Parent Info [Sample Parent](#)
 Sample Tracking [Remaining Samples](#)

Sample Type Information:

Sample Type Name TCL
 Sample Type Description T-cell line (from PBMC's)
 If Active Y
 Sample Type Note NIB Cell Type

Processed Person Information:

Sample Processed by Cimsuser, Ptms
 Phone Number
 Email Address wangg@ninds.nih.gov

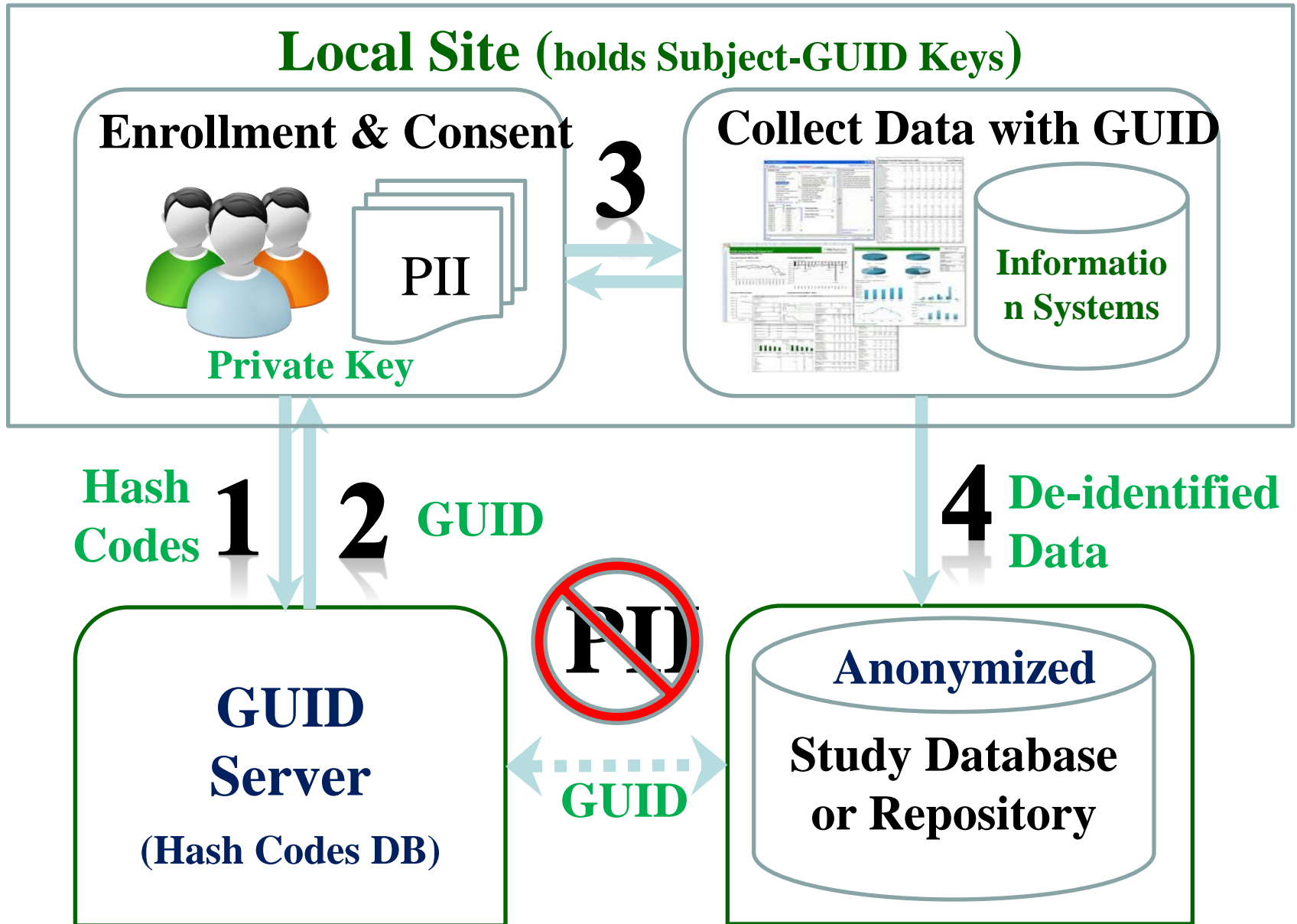
Source/Patient Information:

Source/Patient Name MCGREEVY, Elizabeth (nmn)
 Source/Patient ID in CSIS 558
 Protocols Patient Assigned T-N-0398;

Box1的存放情況

50

如何做嚴謹的匿名化?



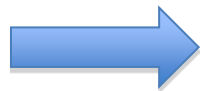
可利用讀卡機收集健保卡中的 人口學資料編GUID碼

- 姓名
- 身份證字號
- 出生年、月、日
- 性別

1 插入健保卡



2



個人資料由醫師自行保管，CSIS中只存放編碼後的資料，以符合GCP的規定



Authentication

Enter 1st Password:

1

Cancel OK



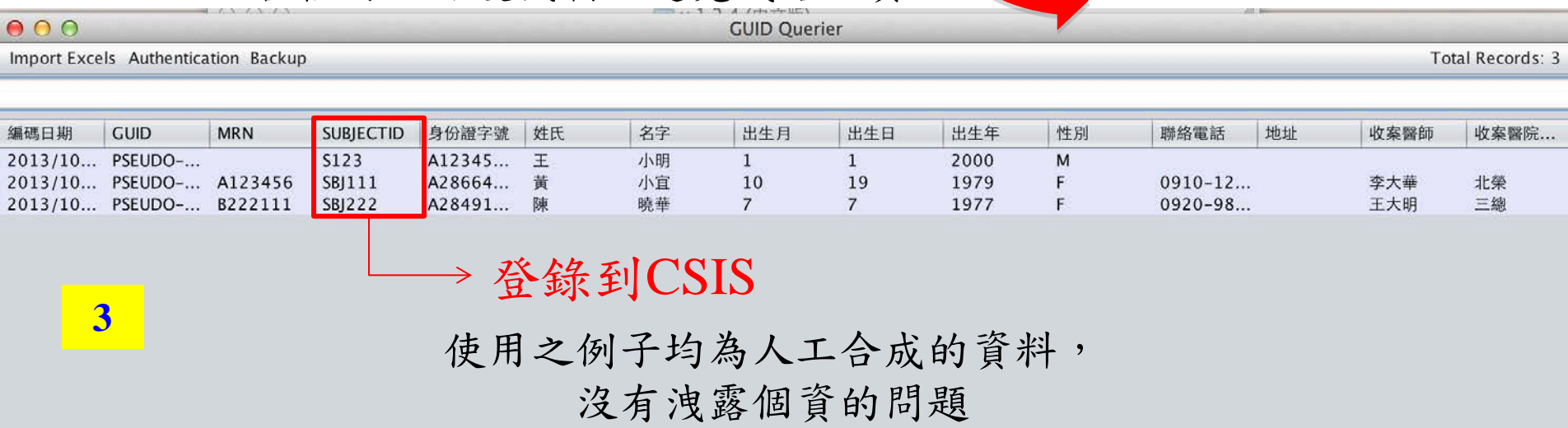
Authentication

Enter 2nd Password:

2

Cancel OK

Querier程式有編輯與查詢功能，並有各種安全防護機制，避免洩露個資



GUID Querier

Import Excels Authentication Backup Total Records: 3

編碼日期	GUID	MRN	SUBJECTID	身份證字號	姓氏	名字	出生月	出生日	出生年	性別	聯絡電話	地址	收案醫師	收案醫院...
2013/10...	PSEUDO-...		S123	A12345...	王	小明	1	1	2000	M				
2013/10...	PSEUDO-...	A123456	SBJ111	A28664...	黃	小宜	10	19	1979	F	0910-12...		李大華	北榮
2013/10...	PSEUDO-...	B222111	SBJ222	A28491...	陳	曉華	7	7	1977	F	0920-98...		王大明	三總

3

→ 登錄到CSIS

使用之例子均為人工合成的資料，沒有洩露個資的問題

虛擬人體生物資料庫採資料集中， 檢體分散的方式運作

臺灣婦女癌症生物資料庫

登入

臺灣婦女癌症臨床—病理組織網路及生物資料庫

「生物資料庫」目的是為了收集人類生物檢體和組織，讓研究者有通路使用這些有價值的檢體，方便做生物與醫學科學方面相關的研究。

了解如何申請檢體

申請帳號

檢體庫存

Log in for users,
research nurses,
managers, PI, Co-PIs,
and user committee



資訊

新聞

簡介

歷年申請名單

論文發表

相關法規

關於申請檢體

表單下載

聯絡我們

相關連結

服務項目

背景，組織與架構，合作醫院，使用者委員會

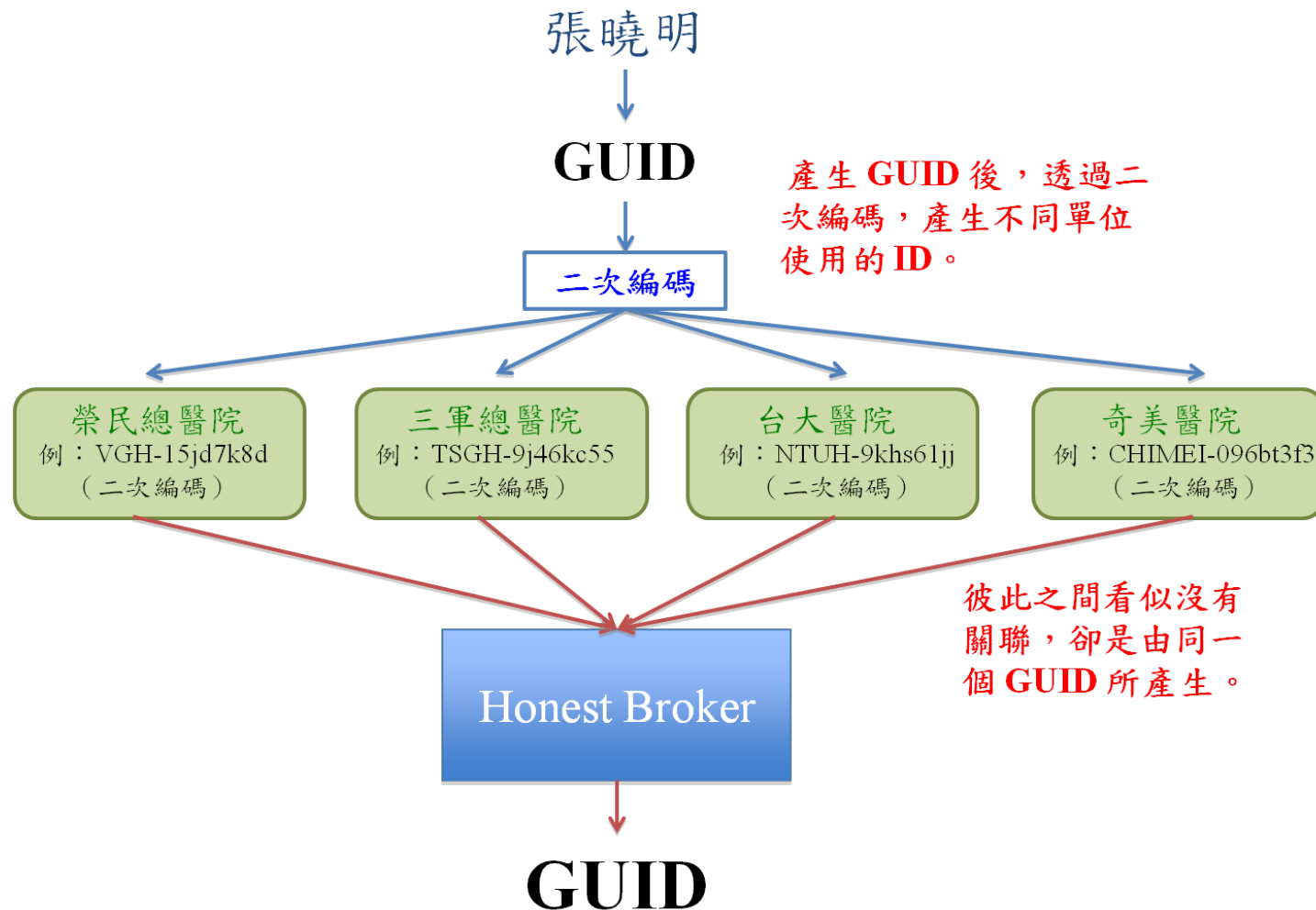
審查原則，申請流程，註冊，查詢申請進度，收費標準，收費方式

檢體申請表，送審文件清單，檢體領取單下載，檢體領取委託書

現有檢體量及類別，臨床、病理及流行病學資料

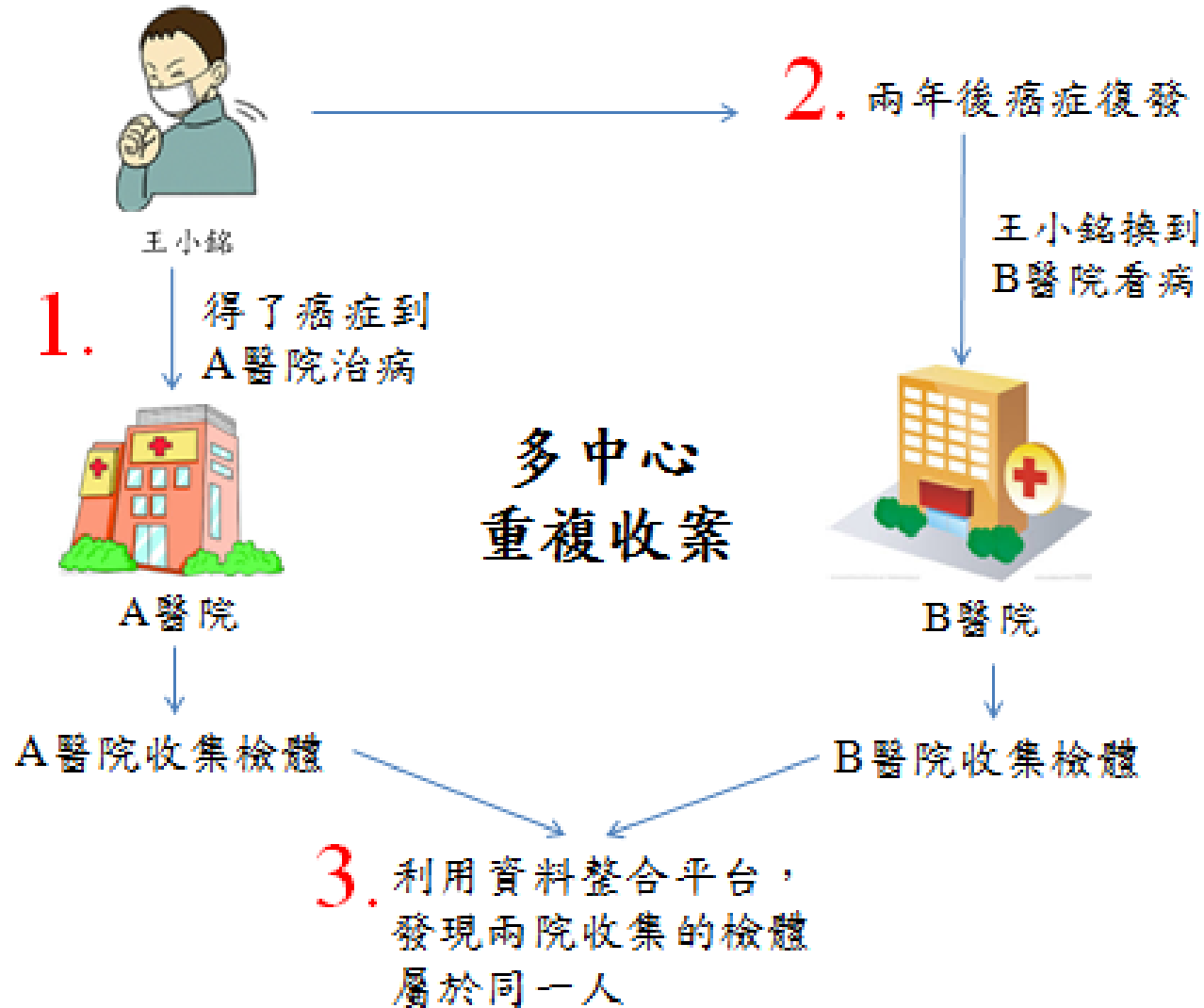


相同的人在不同醫院會編出不同的GUID-T，
在資料匯整時會檢查是否是同一人



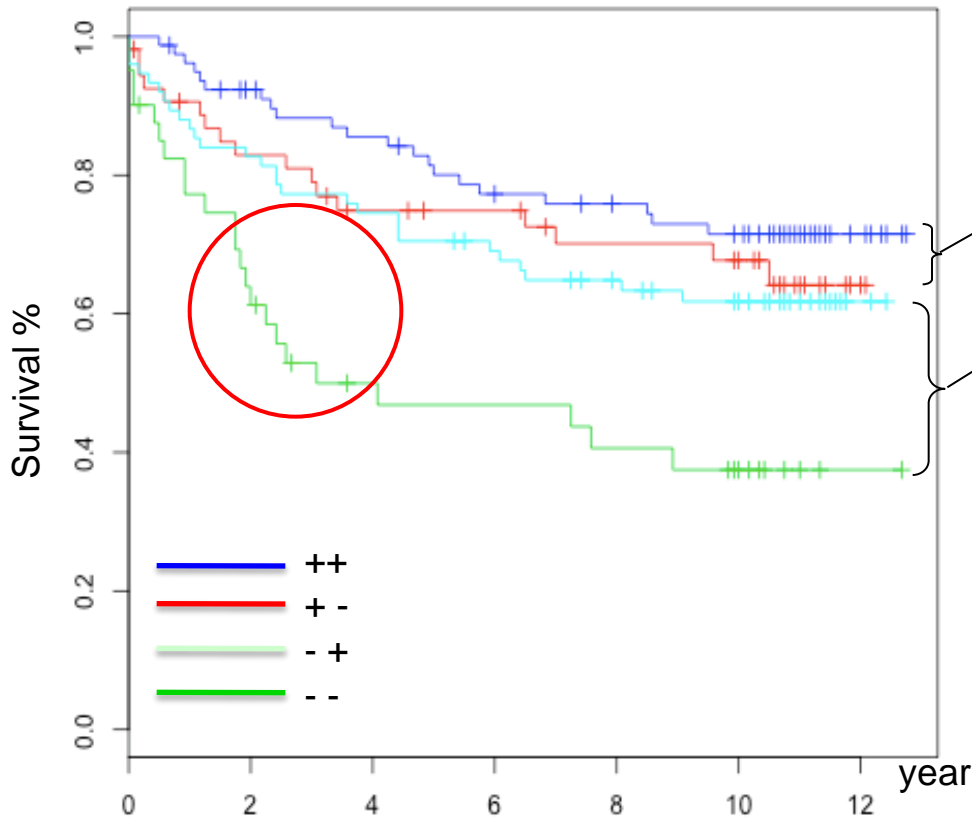
只有 **Honest Broker** 知道還原的演算法，回復原始 **GUID**。

婦癌人體生物資料庫的成功案例

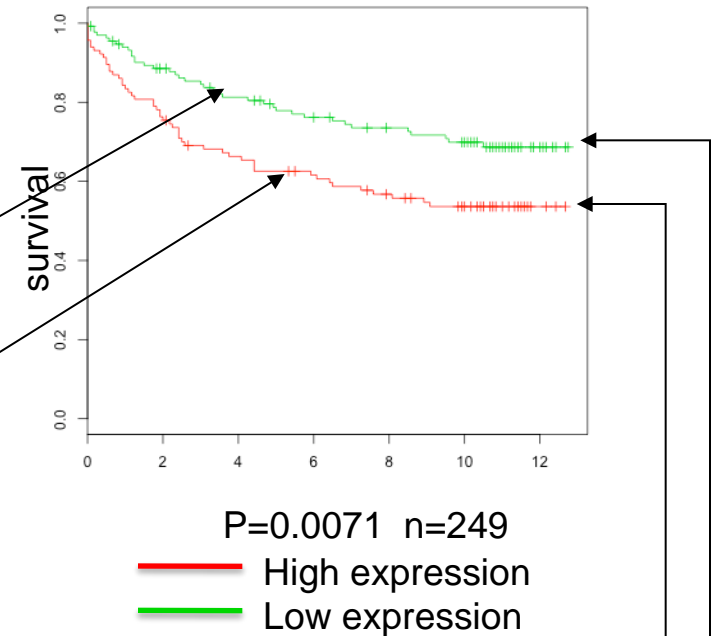


整合基因表現與臨床數據尋找生物標記

Survival curve of biomarker A and biomarker B combinations



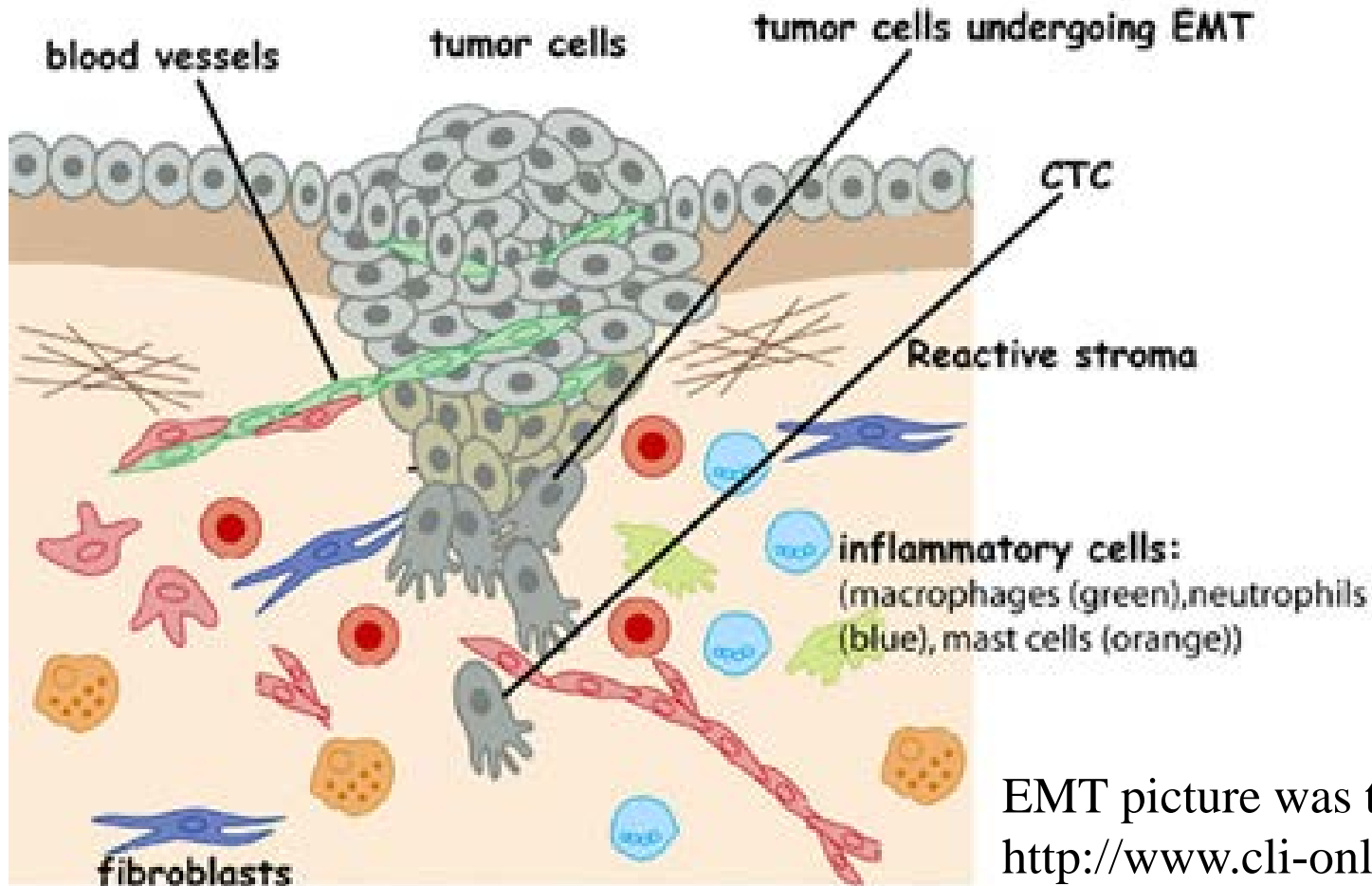
Survival curve of biomarker A or B



	B +	B -
A +	79	75
A -	54	41

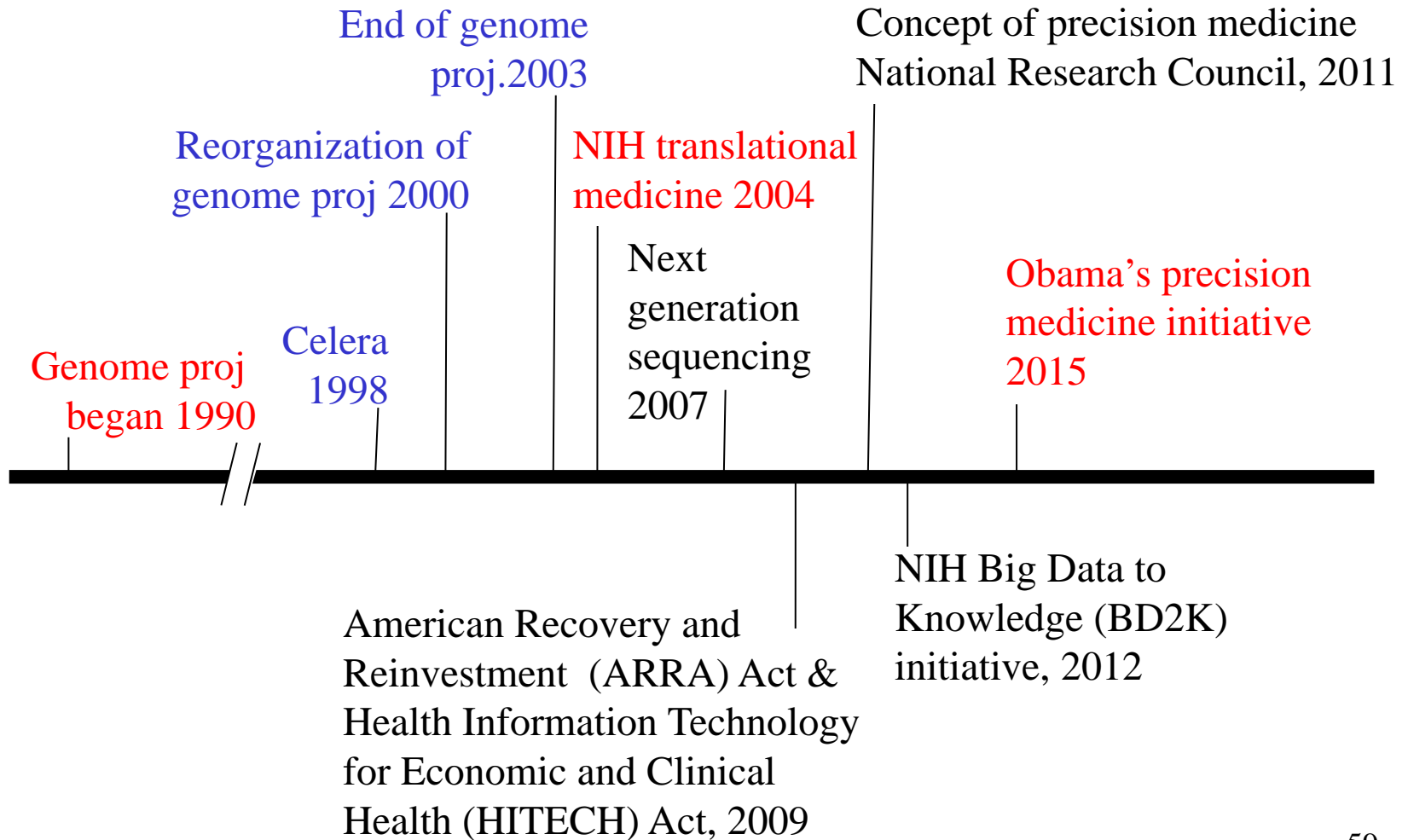
長期追蹤可以顯示趨勢

讓存活率降低的「基因對」可能參與 Epithelial-mesenchymal transition (EMT)



EMT picture was taken from
<http://www.cli-online.com/>

歷史可以幫助我們瞭解未來的趨勢



由大數據到知識(Big Data to Knowledge, BD2K)

UCSD is inviting Taiwan to establish a “data discovery index” jointly

Dear Prof. Yang,

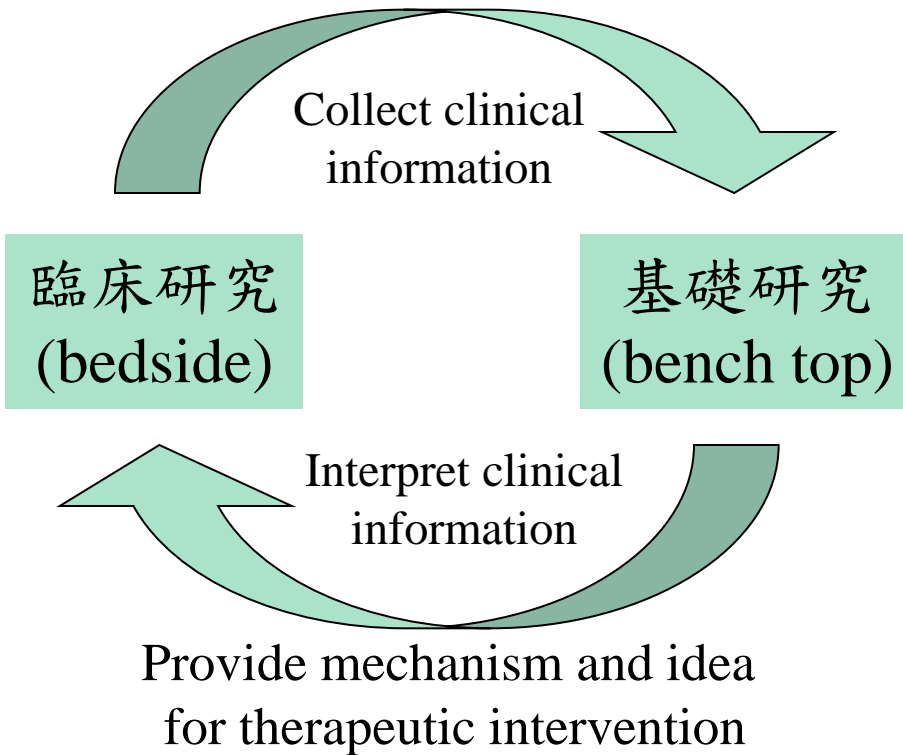
... NIH is calling for the establishment of a consortium to create a "**Data Discovery Index**" to facilitate sharing and reuse of biomedical data.

<http://grants.nih.gov/grants/guide/rfa-files/RFA-HL-14-031.html>

... Please let me know if you are interested. We sincerely hope that you can join this consortium and we think your support is crucial to the approval from NIH. Also, it would be great if you can help us reach out to other Asian countries, like Japan, China, Hong Kong, and South Korea.

將基因體分析的結果 轉化為醫學上的應用

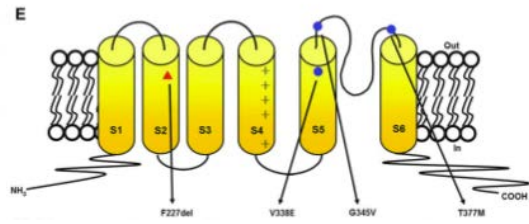
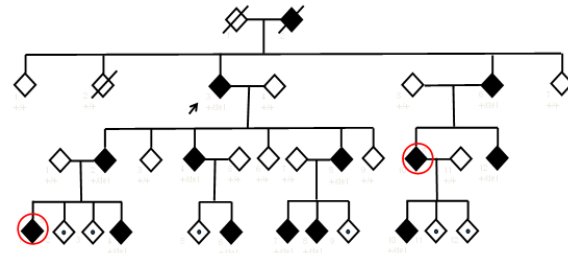
Provide case and control,
clear phenotype definition



利用外顯子(exome)定序， 確認造成疾病的突變位點

第22型脊髓小腦萎縮症

	IV-1	III-10
Total bases sequenced	10,773,734 Kb	10,177,251 Kb
Raw heterozygous variants (number)	165,125	156,440
Common heterozygous variants	35,953	-
Variants not in dbSNP or 1000 genomes database	8,102	-
Variants not found in other 10 non-SCA patients' exome data	5,715	-
Variants resulting in amino acid sequence changes	489	-
Variants present in the linked region (1p21-q23)	11	-
Variants completely segregating with disease phenotype in the pedigree	1	-



Rank	ICN
1	S100A6
2	KCNN3
3	NGF
4	KCNA2
5	KCNC4
6	KCND3
7	KCNA3
8	SPTA1
9	HIST2H2BE
10	SHC1

Nature Reviews Neurology 8, 472 (2012); published online 14 August 2012; doi:10.1038/nrneuro.2012.165

Mutations in *KCND3* linked to spinocerebellar ataxias

The efforts of five independent research groups have identified mutations in the *KCND3* gene that cause spinocerebellar ataxia (SCA) type 19 and type 22. The groups used a combination of genetic linkage analysis in affected family pedigrees and exome sequencing to identify mutations in *KCND3*, which encodes voltage-gated potassium channel K_v4.3 subunit—a protein that is important in membrane repolarization.

“...mutations in *KCND3* ... might underlie ... SCA19 and SCA22”

“We screened through numerous candidate genes without success,” explains Bing-wen Soong, who led one of the teams, “until we performed exome sequencing to pinpoint a defect in ... *KCND3*, which is implicated in cardiac arrhythmia but not known to cause neurological diseases.”

SCAs are a heterogeneous group of neurodegenerative disorders characterized by extrapyramidal and peripheral nervous system involvement and autosomal dominant inheritance. The disease loci for SCA22 and SCA19 were previously

mapped to chromosome 1p21–q21 by linkage analysis in large Chinese and Dutch kindreds, respectively. SCA19 and SCA22 disease loci overlap, and the disease-causing mutations were previously thought to be allelic.

Rare disorders such as the SCAs often occur in small pedigrees, in which traditional genetic analysis cannot identify the genes involved. Margit Burmeister

此突變在多各族裔的病人身上都有觀察到，因此Nature Reviews Neurology特別撰文介紹。

and Soong independently identified mutations in the *KCND3* gene in the original SCA19 and SCA22 families. Burmeister’s group and Giovanni Stevanin’s group also identified new mutations in two other families. Three further mutations in *KCND3* were identified in Japanese SCA22 kindreds by Shoji Tsuji’s group. Notably, none of the mutations identified in SCA19 and SCA22 families was found in healthy controls.

Wild-type K_v4.3 protein is associated with the plasma membrane, whereas the mutated protein found in patients is retained in the endoplasmic reticulum. Electrophysiological studies showed that mutations in *KCND3* impair both channel function and neuronal excitability, suggesting these changes might underlie the slowly progressive cerebellar neurodegeneration observed in SCA19

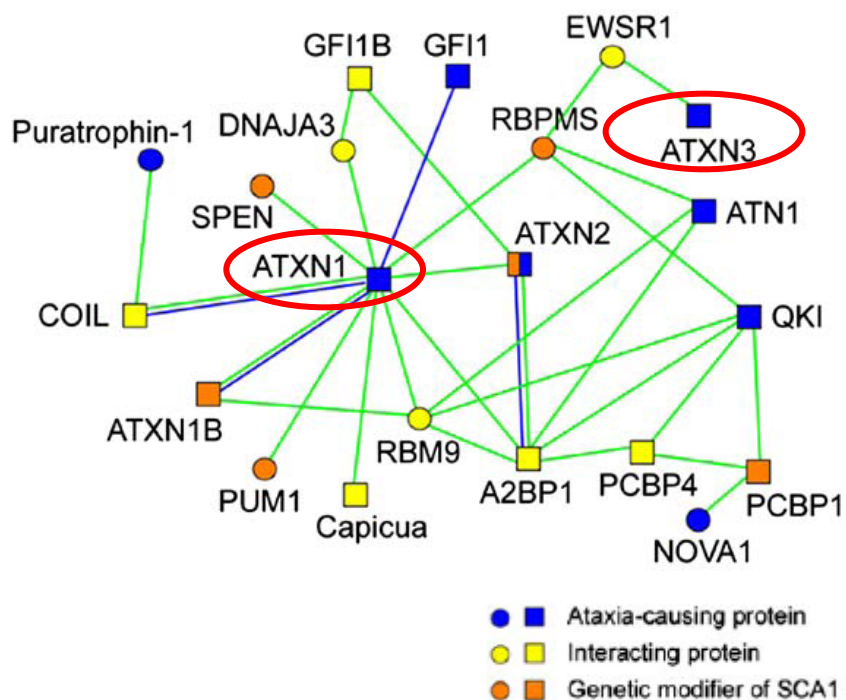
diagnosis of genetic defects in individuals without a family history of SCA—still the majority of patients with ataxia.

Ellen Bible

Original articles Lee, Y.-C. et al. Mutations in *KCND3* cause spinocerebellar ataxia type 22. *Ann. Neurol.* doi:10.1002/ana.23701 | Duarni, A. et al. Mutations in potassium channel *KCND3* cause spinocerebellar ataxia type 19. *Ann. Neurol.* doi:10.1002/ana.23700

網路生物學的啟發：致病基因的預測

Thirty-one legged race



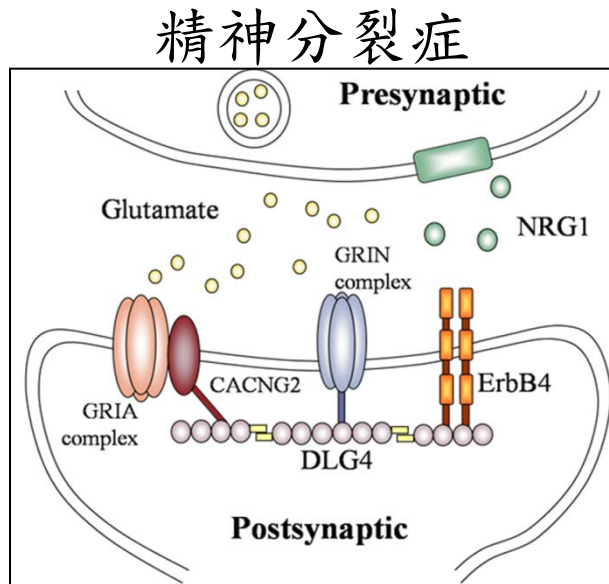
<http://web-japan.org/kidsweb/archives/life/action/06-02/act0602.html>



Pathways /
regulatory circuit

Lim et al. (2006) Cell 125, 801-814

利用病人cohort，預測疾病的形成機制



Recent evidence indicates that enhancing N-methyl-d-aspartate (NMDA) neurotransmission with the treatment of NMDA/glycine site **agonists**, such as **d-serine**, or a glycine transporter-1 (GlyT-1) antagonist, N-methylglycine (**sarcosine**), can improve symptoms of schizophrenia.

...

GlyT-1 inhibitor is more efficacious than the NMDA/glycine site agonist in treatment for schizophrenia, including life quality and global function, at the dosages tested.

Hsu et al (2008) A protein interaction-based model for schizophrenia. BMC Bioinformatics

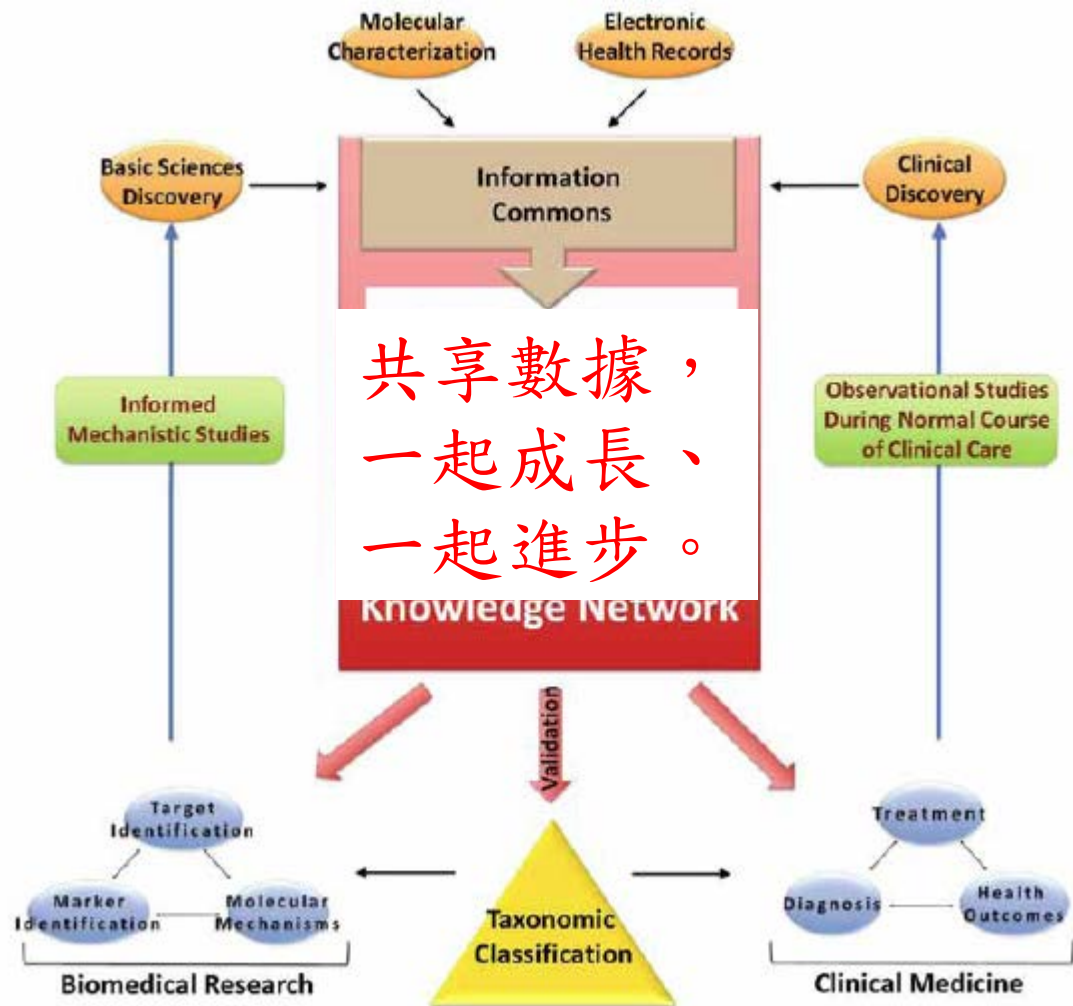
Lane HY, Lin CH, Huang YJ, Liao CH, Chang YC, Tsai GE. (2009) Int J Neuropsychopharmacol. 4, 1-10.

欲達到精準醫學的目標，需要建立知識網路(knowledge network)

Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a **New Taxonomy of Disease**.

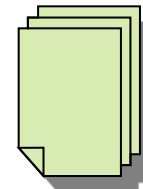
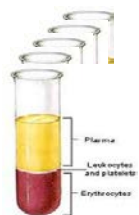
An **expert consensus report** prepared by the Committee on a Framework for Developing a New Taxonomy of Disease

© 2011 The National Academy of Sciences



臨床試驗聯盟

藥物基因體學實驗室 (PGL)



Biospecimen

Biospecimen bank

Sequenom / OncoCarta analysis

Report



Oncocarta Gene output report (Version 1)

ABL1 | AKT1 | AKT2 | BRAF | CDK | EGFR | ERBB2 | EGFR1 | EGFR3 | FLI3 | HRAS | JAK2 |

ABL1 (top)

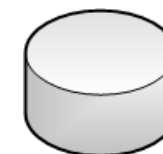
	Assay	Mutation Type	Allele	Mutation Frequency	Mutation
1.	ABL1_1	G250E	A	0.000Low 2-30	WT
2.	ABL1_10	F317L	G	0	WT
3.	ABL1_11	M351T	C	0	WT
4.	ABL1_12	E355G	G	0	WT
5.	ABL1_13	F359V	G	0	WT
6.	ABL1_14	H396R	G	0	WT



Decision making



CSIS



Genotype database

整合基因突變與臨床信息，供醫師決策使用

生物資訊與生物統計中心 (BBC)

美國總統2015宣示推動「精準醫學」

A New Initiative on Precision Medicine

Francis S. Collins, M.D., Ph.D., and Harold Varmus, M.D.

“Tonight, I’m launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier.”

— President Barack Obama, State of the Union Address, January 20, 2015

Collins, F.S. and Varmus, H. (2015) A New Initiative on Precision Medicine. *The New England journal of medicine*.

<http://www.nih.gov/precisionmedicine>

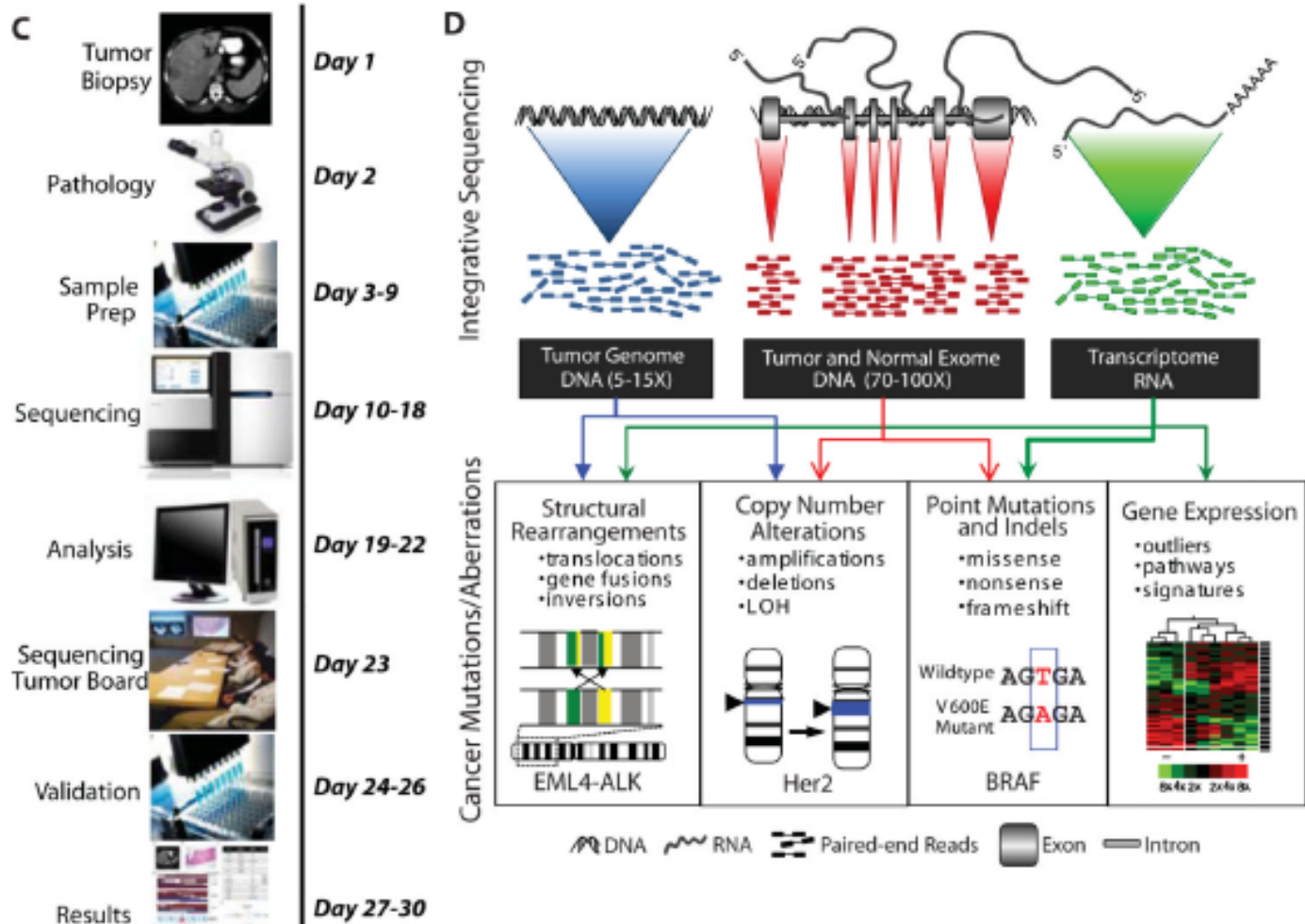
「精準醫學」強調個人化醫療 與轉譯醫學的整合

- Precision medicine is an emerging approach for disease prevention and treatment that takes into account people's **individual** variations in genes, environment, and lifestyle.
- The precision medicine initiative will generate the scientific evidence needed to move the concept of precision medicine into clinical practice.

– Revised from <http://www.nih.gov/precisionmedicine>



以分子層次的證據做癌症治療決策參考

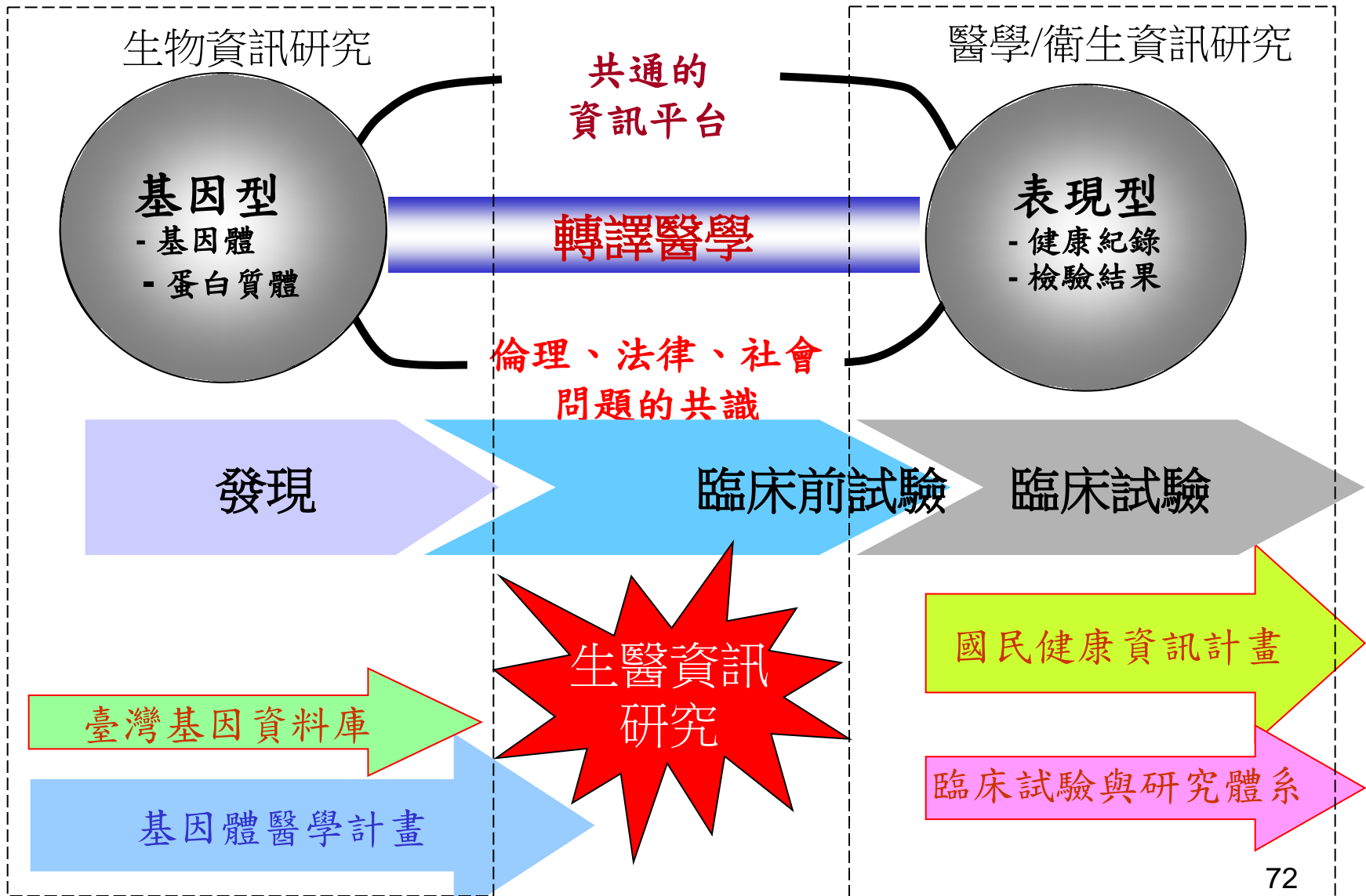


Roychowdhury et al (2011) Personalized Oncology Through Integrative High-Throughput Sequencing: A Pilot Study. *Sci Transl Med.* 3(111): 111ra121

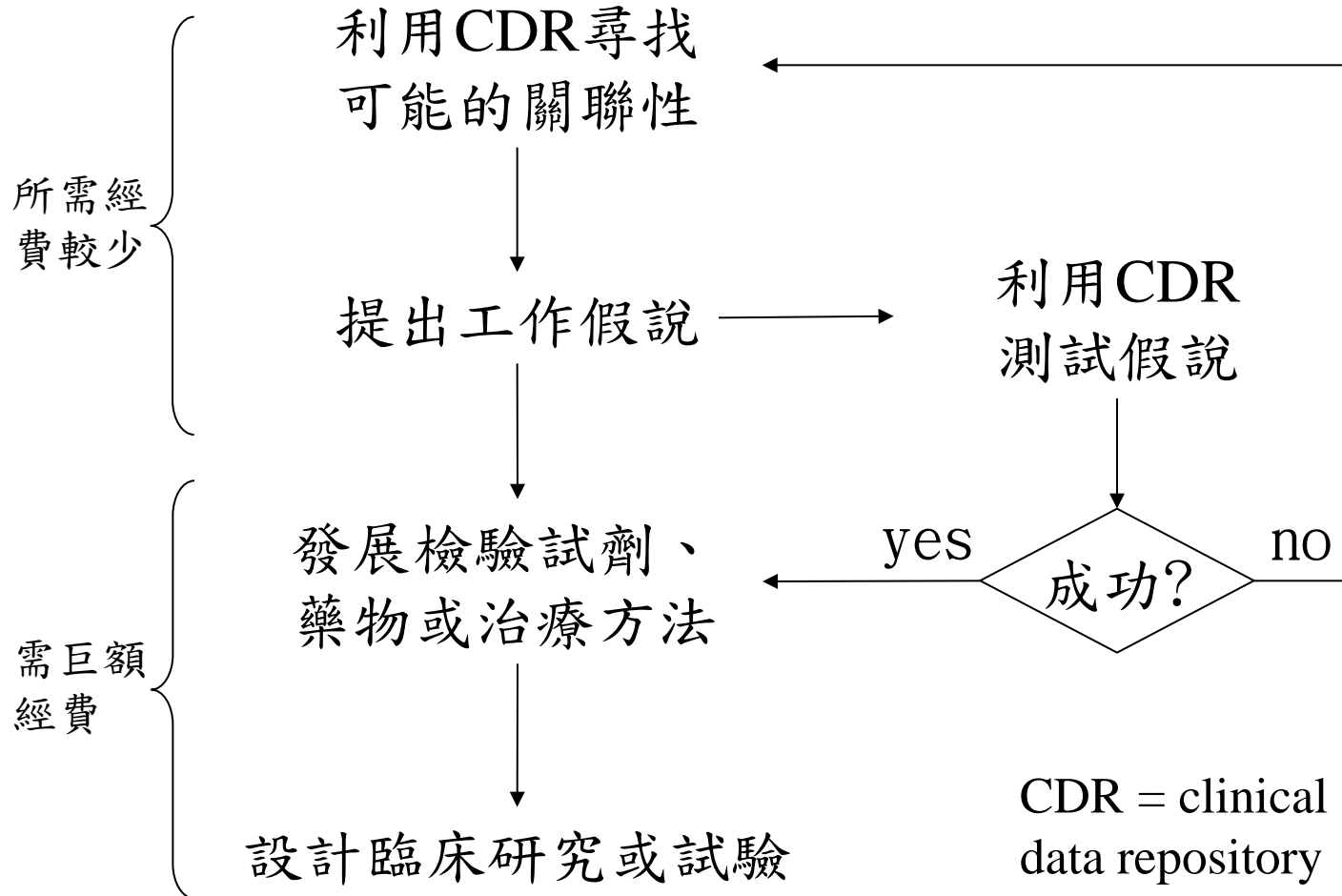
短期與長期目標

- 短期目標 – 強化癌症研究
 - Innovative clinical trials of targeted drugs for adult and pediatric cancers
 - Use of combination therapies
 - Knowledge to overcome drug resistance
- 長期目標 – 收集研究對象群(cohort)
 - Create a research cohort of over 1 million American volunteers who will share genetic data, biological samples, and diet/lifestyle information. All linked to their electronic health records if they choose.
 - Pioneer a new model for doing science that emphasizes engaged participants, responsible data sharing, and privacy protection.

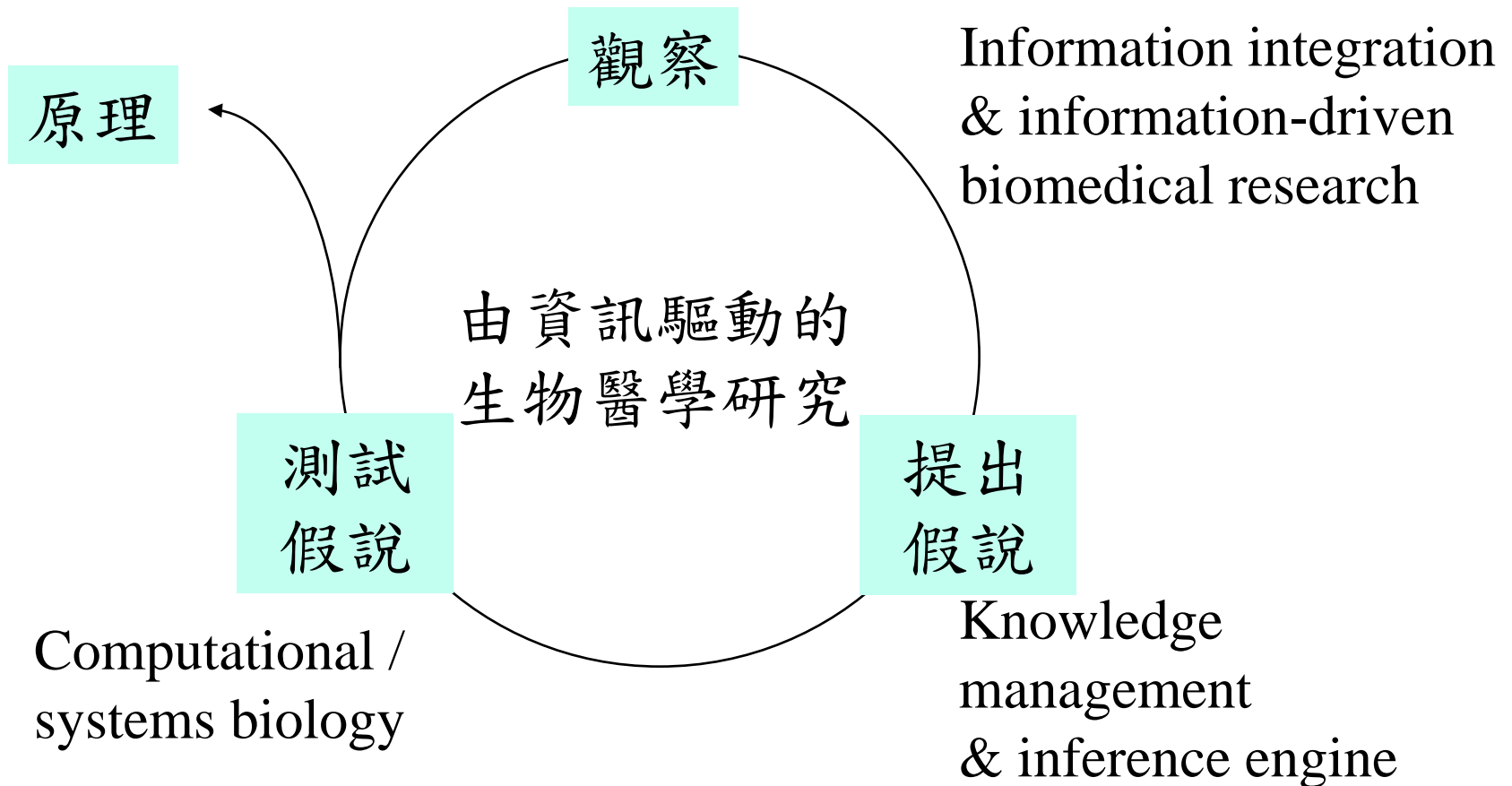
整合生醫資訊， 開創利用大數據的新世代



結合醫療、生活與基因資訊，迎向 智慧醫療的新世代



跨領域是成功的關鍵



Backup slides

研究對象群(cohort)的可能應用

- Advance pharmacogenomics, the right drug for the right patient at the right dose
- Identify new targets for treatment and prevention
- Test whether mobile devices can encourage healthy behaviors
- Lay scientific foundation for precision medicine for many diseases

– Revised from <http://www.nih.gov/precisionmedicine>