AI與大數據相關案件倫理與審查重點

何之行 博士
中央研究院歐美所副研究員
E-mail: chihho@sinica.edu.tw
大林慈濟醫院 IRB課程
15 March 2025

1.2K

TECHNOLOGY NEWS 29 April 2016

Revealed: Google AI has access to huge haul of NHS patient data

A data-sharing agreement obtained by **New Scientist** shows that Google DeepMind's collaboration with the NHS goes far beyond what it has publicly announced

Gathering information
Oli Scarff/AFP/Getty Images

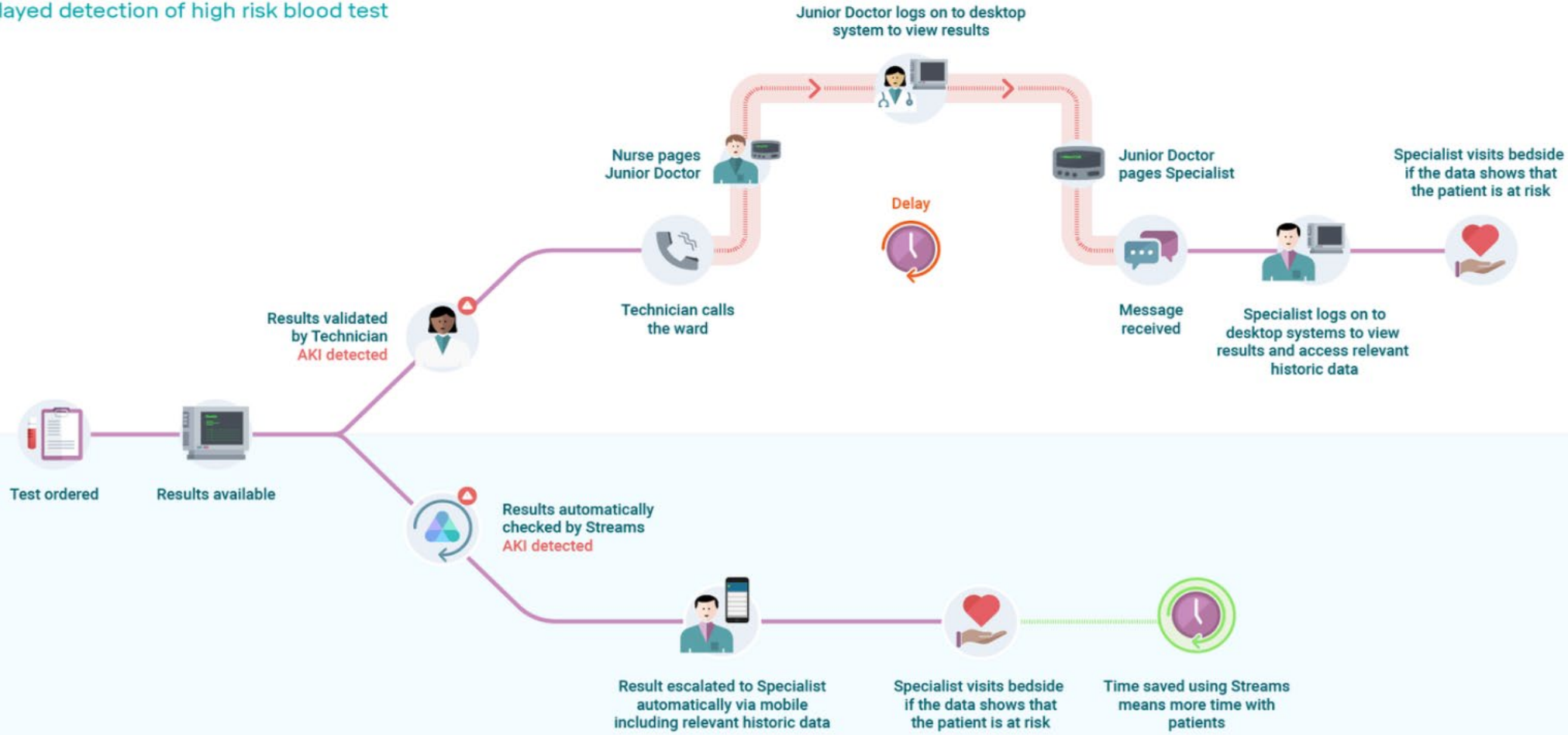
Advertisement

How HIV treatment became global

Sponsored by GILEAD

How we're helping today | D
安全 | https://deepmind.com/applied/deepmind-health/streams/#image-812
Share

Delayed detection of high risk blood test



Accelerated detection and direct notification

ROYAL FREE HOSPITAL- GOOGLE DEEPMIND

- NHS- DeepMind partnership: Information Sharing Agreement, Sep 2015.
- DeepMind would process patients records in the past 5 years with personal identifiable information held by Royal Free to develop a new smartphone app called “Streams”.
- DeepMind had received 1.6 million patient records without patients’ consent
- No prior consulting to relevant regulatory bodies (ICO, HRA)
- UK’s Information Commissioner's Office (ICO) ruled that the Royal free London, National Health Service (NHS) Foundation Trust **failed to** comply with the Data Protection Rule. (July 2017)

Royal Free - Google DeepMind trial failed to comply with data protection law

Date **03 July 2017**

Type **News**

The ICO has ruled the Royal Free NHS Foundation Trust failed to comply with the Data Protection Act when it provided patient details to Google DeepMind.

The Trust provided personal data of around 1.6 million patients as part of a trial to test an alert, diagnosis and detection system for acute kidney injury.

But an ICO investigation found several shortcomings in how the data was handled, including that patients were not adequately informed that their data would be used as part of the test.

INFORMATION COMMISSIONER'S OFFICE (ICO)

- Data transfer without explicit consent
- No privacy impact assessment had been carried out
- The scope of data transfer was too much extended
- Beyond Patients' reasonable expectations on how public data will be managed
- Data Controller (Royal Free) vs Data Processor (DeepMind)
- Transparency

PUBLIC-PRIVATE PARTNERSHIP (PPP) MODEL

- Articulate public benefits
- Ex: making “Streams” app a public property or co-owned by NHS
- Ex: Some of the profits are directed back to the public health system
- Equity and fair distribution of benefits
- Be transparent
- The PPP could consider establishing a joint governance body to guide the data sharing process.

AI AND PRIVACY

- AI/ML: as much training data as possible vs. data minimisation.
- Secondary use of data (資料二次利用)
  - Consent models? (同意範圍)
  - De-identification? (去識別化)
- Commercial Access: PPPs model (商業近用、產學合作)
- Transparency + Accountability ? (透明性、可課責性)
- Privacy Impact Assessment (PIA)-- from data security to fairness ?

個人資料保護法

- 何謂個人資料?

姓名、出生年月日、ID、指紋、職業、婚姻、教育、病歷、醫療、基因、健康檢查、聯絡方式、社會活動 etc. + 其他得以直接或間接方式識別該個人之資料。(Q: 是否具識別該特定個人之可能性?)

- 保護範圍?

目的特定(限縮)原則: 個人資料之蒐集、處理與利用，不得逾越特定目的之必要範圍 (§5)。

- 區分敏感性個資與非敏感性個資
- 敏感性個資 (醫療、基因、健康檢查) : 原則不得蒐集處理或利用，例外允許 (§6)。

敏感性個資之蒐集處理或利用(§6但)

- 法律明文規定。
- 公務機關執行法定職務或非公務機關履行法定義務必要範圍內，且事前或事後有適當安全維護措施。
- 當事人自行公開或其他已合法公開之個人資料。
- 公務機關或**學術**研究機構基於醫療、衛生或犯罪預防之目的，為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人。 **(Q: 業界? 去識別化標準?)**
- 為協助公務機關執行法定職務或非公務機關履行法定義務必要範圍內，且事前或事後有適當安全維護措施。
- 經當事人書面同意。但逾越特定目的之必要範圍或其他法律另有限制不得僅依當事人書面同意蒐集、處理或利用，或其同意違反其意願者，不在此限。

DE-IDENTIFICATION

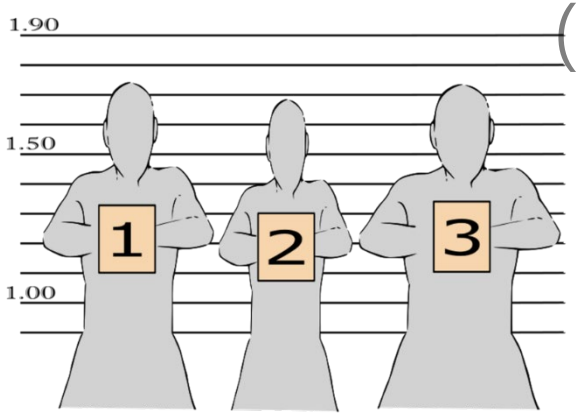
- 去識別、無從識別、去連結 vs. 匿名化 (Q: 是否無回溯可能性?)
- Big data challenges..
- USA: HIPAA (Health Insurance Portability and Accountability Act)
  - removal of 18 identifiers: deemed to be not personal data
- EU GDPR: Pseudonymised (假名化) data: still personal data
  - NOT a valid legal basis for processing data
- TW: standard not clear– suggestion: privacy impact assessment (PIA) + Data Access Committee

Pseudonymization and anonymization
假名化及匿名化

They are just safeguards (安全維護措施) and conditions for processing,
not a legal base (非為處理資料之合法充足條件)

Pseudonymization (假名化，去識別):

personal data can no longer be attributed to a specific data subject without the use of additional information (e.g. a key or encryption code)
(如掌握特定資訊，如金鑰，則仍可連結)



Anonymization (匿名化，去連結):

the data subject no longer identifiable (無從識別)

人體研究法

- 人體研究：指從事取得、調查、分析、運用人體檢體或個人之生物行為、生理、心理、遺傳、醫學等有關資訊之研究。(第四條)
- 去連結：指將研究對象之人體檢體、自然人資料及其他有關之資料、資訊（以下簡稱研究材料）編碼或以其他方式處理後，使其與可供辨識研究對象之個人資料、資訊，永久不能以任何方式連結、比對之作業。(第四條)

人體研究法 (第十四條)

- 研究材料於**研究結束**或第十四條第一項第八款所定之保存期限屆至後，應即銷毀。**但**經當事人同意，或已去連結者，不在此限。
- 使用**未**去連結之研究材料，逾越原應以書面同意使用範圍時，應再依第五條、第十二條至第十五條規定，辦理審查及完成告知、取得同意之程序。
- 未去連結之研究材料提供國外特定研究使用時，除應告知研究對象及取得其書面同意外，並應由國外研究執行機構檢具可確保遵行我國相關規定及研究材料使用範圍之擔保書，報請審查會審查通過後，經主管機關核准，始得為之。

發文單位：法務部

發文字號：法律 字第 10703512280 號

發文日期：民國 107 年 09 月 04 日

資料來源：法務部法規諮詢意見

相關法條：個人資料保護法 第 2、6、15、19 條（104.12.30）

要　　旨：公務機關將所收載醫療影像資料上之特種個人資料進行去識別化處理如符合個人資料保護法第 6 條第 1 項但書之規定，應可認並未逾越原先蒐集之特定目的，而得依據原先蒐集時之同一合法事由為之，而**去識別化**後如可認非屬個人資料，則後續提供外界利用，自無個人資料保護法適用

主　旨：有關貴署參與國家發展委員會推動之「亞洲．矽谷試驗場域計畫」，**規劃執行健保醫療影像倉儲建置與人工智慧應用，未來擬開放去識別化之醫療影像資料供外界應用**，其中涉及個人資料保護法之解釋適用疑義乙案，復如說明二，請查照參考。

人權業務

- 人權議題報告
- 參與憲法法庭
- 系統性訪查研究
- 人權教育推廣
- 國際合作交流
- 國內合作交流

首頁 › 人權業務 › 參與憲法法庭

參與憲法法庭

111年憲判字第13號：健保資料庫案

類別：言詞辯論意見書，提供書面意見

本案係蔡季勳、邱伊翎、施逸翔、滕西華、黃淑英、劉怡顯、洪芳婷等 7 人，就個人資料保護法第6條第1項第4款、個資法第16條第5款之規定之合憲性，聲請釋憲。

本會應司法院邀請，於111年4月26日由王幼玲委員以鑑定機關身分參與言詞辯論。

王幼玲委員表示，國民健康與個人隱私並非全有全無之零和賽局，個人健康資料攸關醫師診斷正確性及疾病防治，相較其他隱私更具敏感性，當代公衛之進步，高度仰賴健康個資之蒐集、處理與利用，且統計及研究結果對國民健康有重要價值，但即便是追求公益，仍應尊重當事人對個資保護之實質權利，並提供適當且具體之保護措施。健保資料庫目前之法制設計，法律授權不夠明確、去識別化要求不夠清楚、相關資訊安全與資料是否符合執行法定職務之最小必要原則，均缺乏足夠之檢驗機制，縱行政機關以相關命令降低隱私外洩之風險，但法制架構不足、隱私保障有限之情形，完全禁止個資當事人享有退出權，違反比例原則。

111年8月12日憲法法庭作出111年憲判字第13號判決。

相關檔案

111年憲判字第13號：健保資料庫案言詞辯論意見書 PDF

111憲判字第13號判決

- 個資法第六條第一項但書第四款：(V)
- 健保法 79、80條: 欠缺健保資料提供目的外利用之明確機制：(X)
- 個資法欠缺獨立監督機制，有違憲之虞。
- 三年內立法或修法補正: 若逾其未完成，當事人可請求停止利用
- 未立即宣告禁制資料釋出
- 隱私保護 vs 研究之公益性: 平衡

111憲判字第13號判決 II

- 個人健康保險資料得由中央健康保險署以資料庫儲存、處理、對外傳輸及對外提供利用之主體、目的、要件、範圍及方式暨相關組織上及程序上之監督防護機制等重要事項等，因為欠缺明確規定，不符法律保留原則而違憲。
- 衛生福利部中央健康保險署就個人健康保險資料之提供公務機關或學術研究機構於原始蒐集目的外利用，由相關法制整體觀察，欠缺當事人得請求停止利用之相關規定而違憲等。

111憲判字第13號判決

- 適用範圍??
- 處理問題: 個人健保資料二次利用有無適法基礎?
- **Q: 個人敏感個資原始蒐集目的外之利用，規範基礎是否完備?**
- 是否具備: 法律明文規定 或 當事人同意?
- 同意之範圍
- 去識別化之限制
- 當事人停止利用

Your Data Matters to the NHS

Information about your health and care helps us to improve your individual care, speed up diagnosis, plan your local services and research new treatments.

In May 2018, the strict rules about how this data can and cannot be used were strengthened. The NHS is committed to keeping patient information safe and always being clear about how it is used.

You can choose whether your confidential patient information is used for research and planning. To find out more visit: nhs.uk/your-nhs-data-matters or download a copy of the patient leaflet.

Our Trust is compliant with the National Data Opt-Out policy and operational guidance.

We follow national and local guidance.

National Library of Medicine
National Center for Biotechnology Information

PMC PubMed Central®

Search PMC Full-Text Archive

Search

Journal List > J Public Health (Oxf) > PMC6306093

J Public Health (Oxf). 2018 Dec; 40(4): e594–e600.
Published online 2018 Mar 26. doi: 10.1093/pubmed/fdy059

PMCID: PMC6306093
PMID: 29590471

The challenge of opt-outs from NHS data: a small-area perspective

Frédéric B Piel, Lecturer in Epidemiology, Brandon L Parkes, Research Associate in Biostatistics, Hima Daby, Data and Information Services Manager, Anna L Hansell, Reader in Environmental Epidemiology, Assistant Director of SAHSU, and Paul Elliott, Chair in Epidemiology and Public Health Medicine, Director of SAHSU

▸ Author information ▸ Article notes ▸ Copyright and License information Disclaimer

This article has been cited by other articles in PMC.

Introduction

Go to: ▸

One of the founding principles of the NHS is that it offers comprehensive, universal and free public health services at the point of delivery. As a result, NHS data provide a huge and invaluable resource of routinely collected primary (e.g. visits to GP practices) and secondary (e.g. hospital admissions,

J Public Health (Oxf)

EU GDPR ARTICLE 6 (4): 資料二次利用之兼容性判斷

- Any link between the original and new purpose (原始目的與進階目的之連結)
- The context in which the personal data have been collected (蒐集資料之背景，考量資料主體與資料控制者間的關係)
- Nature of the personal data, in particular whether special categories of (sensitive) personal data are processed (資料性質；是否為敏感資料)
- The possible consequences of the intended further processing (進階處理可能對資料主體造成之後果)
- The existence of appropriate safeguards, which may include encryption or pseudonymisation. (適當保護措施，如加密或假名化)

大數據、資料探勘、研究申請
(實作規範)

NTUHREC Version：AF-158/03.0

國立臺灣大學醫學院附設醫院研究倫理委員會
使用大量病歷資料進行研究申請表

1. **下述研究若申請免除知情同意，請填寫此表：**
   大數據、資料探勘、人工智慧、建立個別研究用資料庫等類別或申請病歷[註]個案數 1,000 以上。
   註：前述所指之病歷資料，包含健檢與影像學檢查等個人健康與病歷資料。
2. 審查方式標準：
   ＊申請病歷個案數 1,000 以上（資料僅在本院之臺大醫療資料分析專區分析之計畫則為 10,000 筆以上）若申請簡易審查，須符合以下全部條件：
   a.取得之資料須為完全去識別化資料
   b.申請病歷資料的項目有特定範圍
   c.資料僅在院內使用
   d.研究者不含院外人員
   ＊以上若有任一不符，則為一般審查。

一、基本資料			
計畫名稱		**計畫主持人**	
所屬機構/科部			
題目	**說明**		
	☐主持人自行發起： ☐無院外合作者（研究團隊人員均領有本院員工識別證）。 ☐有院外合作者：☐本校系院所之師生；☐校外學術研究機構：＿＿＿＿＿＿＿＿；		

4.資料是否攜出院外 <註：若經審查同意，亦須經本會簽核院方同意後方得攜出，且攜出之資料必須經處理後無可辨識個人資料>	□否，資料僅會在院內使用 □是，資料須攜出院外，否則計畫無法執行，請說明： (1) 必須攜出之理由： (2) 攜出方式：
5.資料保存地點及設備擁有者	※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第___頁。

5-1	資料保存地點	
	設備擁有者（電腦含伺服器）	
5-2	資料運算分析地點	
	設備擁有者（電腦含伺服器）	
5-3	以上設備是否會與外部網路連結或傳輸？	□否，不會與外部網路連結或傳輸。 □是，請說明確保資訊安全之措施： _____。

<註：資料若僅於【臺大醫療資料分析專區】使用，請於計畫書中敘明「向醫療整合資料庫申請之資料僅能於臺大醫療資料分析專區使用」；若非前述狀況或有必要理由，請於計畫書中敘明「遵循資料不得攜出院外」。>

6.本研究使用之資訊科技(IT)		
6-1	資訊科技	□無使用資訊科技，請說明使用之方式或統計工具：_____。 □使用資訊科技，如大數據資料探勘、人工智慧學習運算等，請說明提供者：_____。 【例如：XXXX大學，OOO教授實驗室；OOOO公司...等】
6-2	雲端運算服務	□無使用雲端運算服務； □使用雲端運算服務，由誰提供請說明：□同上 6-1□其他： _____。
7.資料安全及隱私個資保護措施	【保護措施舉例：資料保存於上鎖之OOO研究室，保存資料之電腦與伺服器以密碼保護，資料將於院內保存與運算分析。存取資料之權限管制如9.1.3.1，可識別身分的個資將以代碼或加密保護。由主持人親自確認所有研究人員(含主持人)完成病歷隱私保護與資訊安全訓練、簽署保密聲明切結書、合作研究合約等。】	
	※資料安全及隱私個資保護措施需一併呈現於計畫書中，請說明記載於計畫書第___頁。請說明：	

8.資料保存與結束後處理	※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第___頁。	
8-1	資料保存與期限	□研究結束後_____年銷毀資料 □永久保存，請說明必要理由_____。
8-2	最終處理措施 【例如：儲存資料之硬碟將會	

9.資料類型 **(可複選)**	☐9.1.研究者取得的資料含有可識別身分之資料，研究者不會將之完全去識別化。 ☐9.2.研究者取得的資料含有可識別身分之資料，但研究者會將之完全去識別化。[註] ☐9.3.研究者取得的資料為完全去識別化之資料。 <註：個人資料包括可直接或間接識別身分之個資(identifiers)及其他個資。前者包括最末頁註解之個資項目等。完全去識別化是指資料中不包含可直接或間接識別身分之個資，且於該研究案之研究者無法從取得之資料識別身分。>

二、個資隱私保護(請依 9.資料類型 勾選的類別填寫)

◎勾選「9.1.研究者取得的資料含有可識別身分之資料，研究者不會將之完全去識別化」者，請續完成所有 9.1 之子項：

※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第___頁。

9.1 9.1.1 研究需使用可識別身分個資，否則計畫無法執行。

請說明必要理由：

9.1.2 研究需申請免除知情同意，否則計畫無法執行。

請說明必要理由：

9.1.3 本研究的風險已降至最小

9.1.3.1 有適當的規劃，保護可識別身分的個資，避免不當的使用與揭露？

將接觸病歷資料之研究人員名單：<註：欄位可請自行增列>

姓名	單位	職稱	病歷隱私個資保護訓練課程的完訓日期	使用資料之權限為何？ 【例如：負責資料統計、負責資訊程式撰寫...】
本院（領有本院員工識別證者）				
本校其他系院所之師生				

三、本院與本院研究人員的權益

10.是否有院外人員使用本院病歷資料？【例如：本校系院所之師生、校外學研機構或廠商，或使用廠商提供的雲端運算服務。】

□否

□是，請續填10.(1)~(3)。

10 (1)研究成果的歸屬：研發成果商品化時智慧財產權益之歸屬？

請說明：

_____。

(2)研究成果回饋方式為何？

請說明：

_____。

(3)是否已簽署合作研究合約或產學合作研究合約，約定智財權益的分配及成果保密義務，其相關約款為何？

□是，已簽署合作研究合約或產學合作研究合約，請檢附文件，相關約款記載於第___頁。

□否，預計何時簽署合約或其他說明:_____

申請人簽章 計畫主持人：單位主管：

THANK YOU FOR YOUR ATTENTION