Balancing privacy and best price
in the age of US pension risk transfer

Danielle Prado
Panelist

Senior Associate,
MBM Commercial
LLP

Tom Jones
Panelist

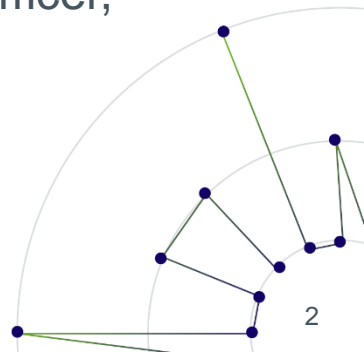Pricing Actuary,
Fidelity & Guaranty
Life Insurance
Company

Brad Malin
Panelist

Accenture Professor
of Biomedical
Informatics,
Biostatics and CD,
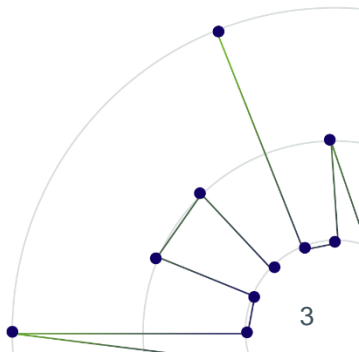Vanderbilt University

Douglas Anderson
Webinar chair

Chief Visionary Officer,
Club Vita

Three perspectives

1. Lawyer: how are US attitudes to data privacy changing?

2. Underwriter: is providing personal identifiers worth it?

3. Data scientist: how can data provision be performed safely?

9-digit vs 5-digit ZIP codes

79936 El Paso, TX
Population 115k

90011 Los Angeles, CA
Population 106k

60629 Chicago, IL
Population 105k

90650 Norwalk, CA
Population 105k

90201 Bell Gardens, CA
Population 101k

On average, a 5 digit ZIP covers 7,800 people – but in many cover more than 100,000 residents

Around 1/3 of 5 digit ZIPs in the US contains at least six of Club Vita's 9-digit longevity groups

9-digit ZIP codes are needed to capture the diversity within each local area

CLUB VITA

1 Changing US data privacy requirements
Danielle Prado

Agenda

- Introduction
- Attitudes to Privacy in the U.S.
- General Regulatory Landscape
- Looking Forward

Introduction

Danielle Prado

Senior Associate

Dual Qualified: Texas and England

E: Danielle.prado@mbmcommercial.co.uk

Tracey Ginn

Partner

Dual Qualified: California and Scotland

Alexander Lamley

Director

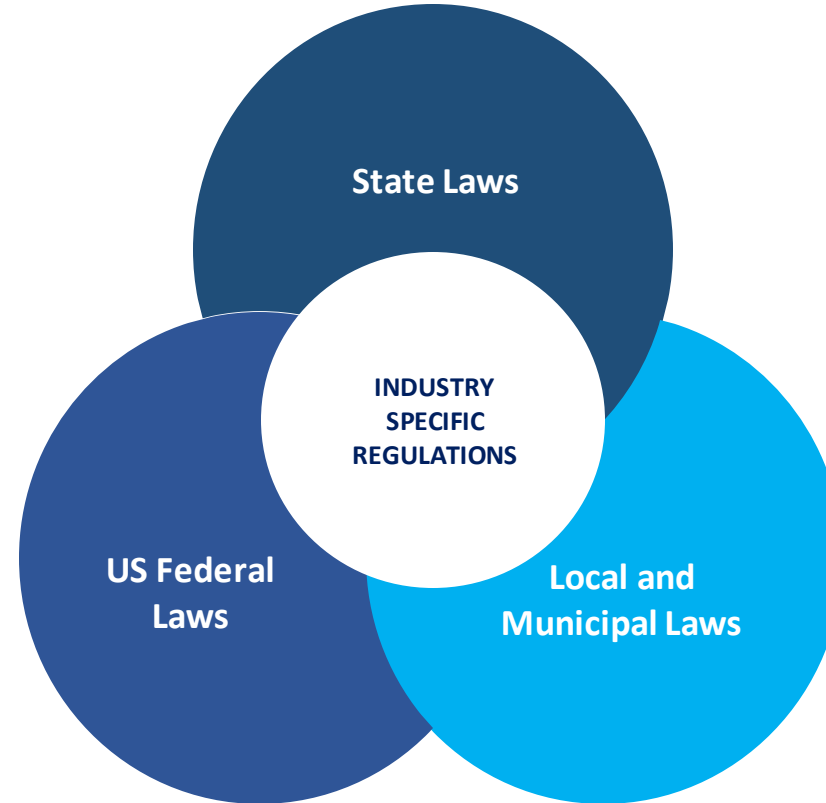Dual Qualified: New York and Scotland

ENTREPRENEURIAL BUSINESS LAWYERS
mbmcommercial.co.uk

MBM

Attitudes to Privacy

- **Europe**
  - fundamental human right
  - Supranational/harmonized

- **United States**
  - integrity of data important but primarily for commercial exploitation
    - privacy never even appears in the Constitution
  - Bottom-up approach – states versus federal government
  - Price to pay for perceived benefits
  - Economic power of big-tech

ENTREPRENEURIAL BUSINESS LAWYERS
mbmcommercial.co.uk

MBM

General Regulatory Landscape

State Laws

INDUSTRY SPECIFIC REGULATIONS

US Federal Laws

Local and Municipal Laws

General Regulatory Landscape

US State Privacy Legislation Tracker 2022

ENTREPRENEURIAL BUSINESS LAWYERS
mbmcommercial.co.uk
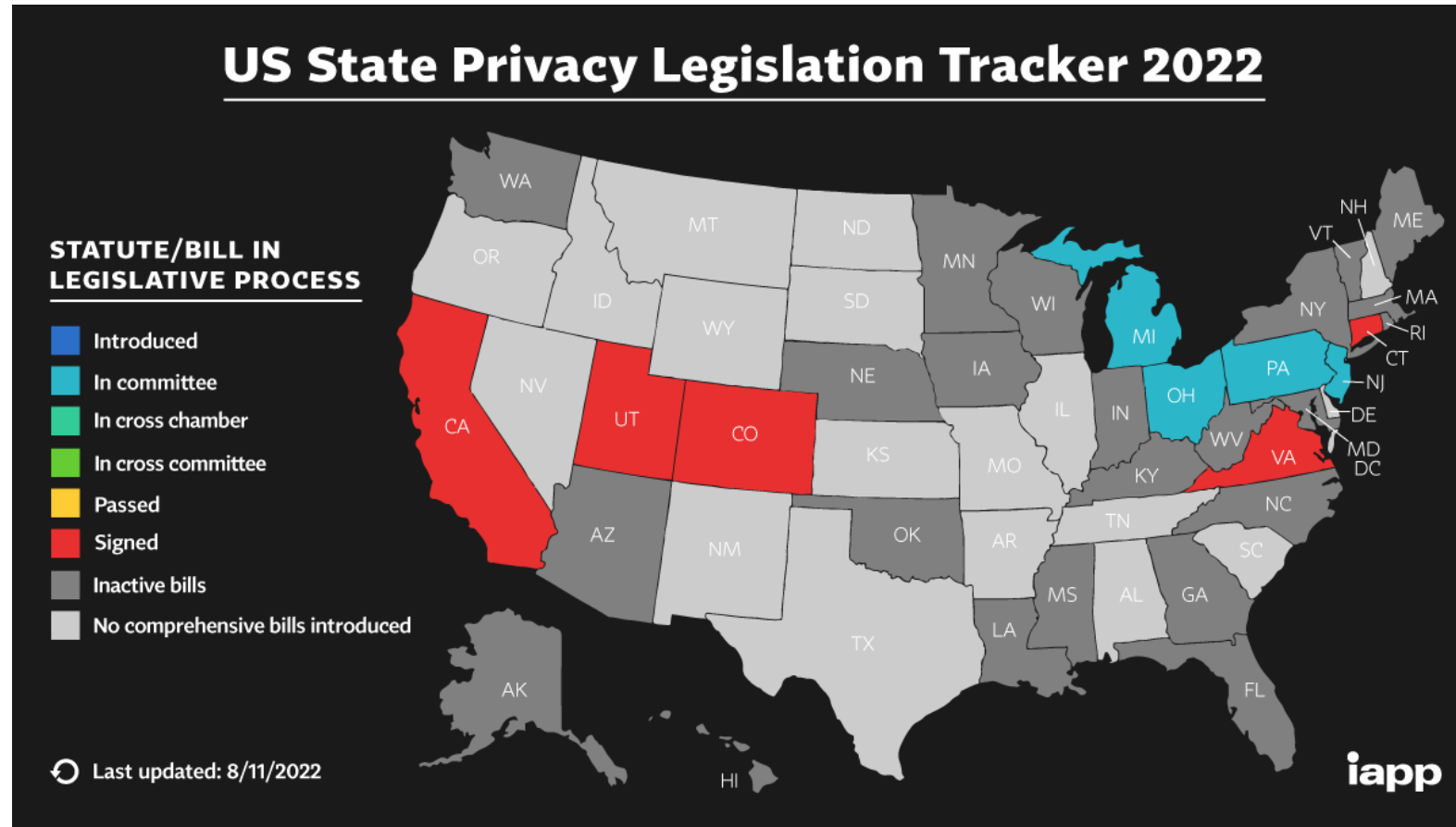
Looking Forward

- **American Data Privacy and Protection Act**
  - First American consumer privacy bill to make it to a full chamber vote in either House or Senate
  - GDPR on a continuum (small data holders and large data holders)
    - Midterm elections, preemption concerns

- **Trans-Atlantic Data Privacy Framework**
  - Executive action (perhaps on 3 October)
  - To address gap after Schrems decision nullifying Privacy Shield
  - Ratification required by European Commission (could take up to 6 months)

2 Is providing personal identifiers justified?
Tom Jones

PRT Mortality Underwriting

▶ PRT transaction pensioner census data file

▶ Insurance company in-force annuity block data

▶ Industry data (e.g. SOA tables)

▶ Client specific Mortality Experience Data (MED)

Mortality underwriting refinements can be worth 1% - 4+% of a PRT transaction price

Market Segment	Deal Size	Data Availability		
		MED	Partial Zip Code (5 digit only)	Full Zip Code (+4 digits)
Jumbo	>$1B	Always	Often	Inconsistent
Large	$500M - $1B	Often	Often	Inconsistent
Mid	$250M - $500M	Sometimes	Often	Inconsistent
Small	<$250M	Rarely	Often	Inconsistent

PRT Transaction Bid Data

- ► Benefit size
- ► Annuitant gender
- ► Plan sponsor industry
- ► Hourly/Salaried
- ► Occupation/work type
- ► Geographic location (state and zip code)
  - ► Full address information is typically excluded from bid data
  - ► 9 digit zip code is rarely available for all records within pension administration database

> Availability and quality of information may drive increased insurer participation and helps eliminate pricing conservatism

Evolution of Zip Code Data Availability Within PRT Transaction Bid Process

2010	2015	2020	2025???
• >$3B sales	• $10B - $15B sales	• $20B - $30B sales	• $35B - $50B sales
• 6-8 issuers	• 10-12 issuers	• 15- 18 issuers	• 20+ issuers
• No Zip data	• 5 digit Zip only	• Inconsistent 9 digit Zip	• Consistent 9 digit Zip

Range of period life expectancy in 9-digit longevity groups

Crude life expectancy at age 65 split by lifestyle groups colored by 9-digit ZIP longevity groups
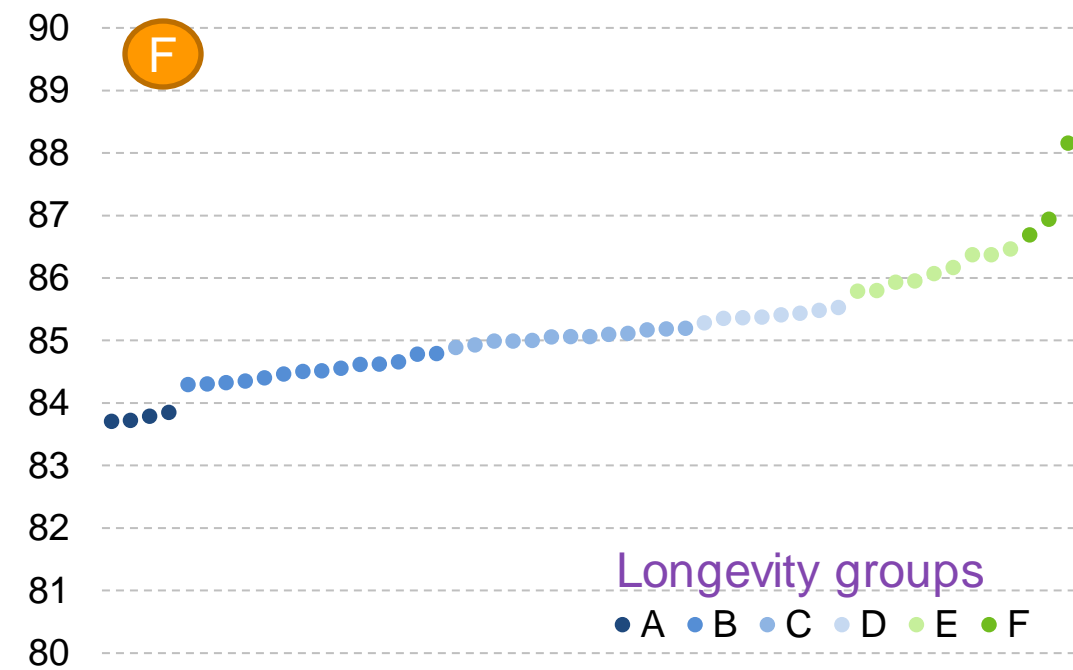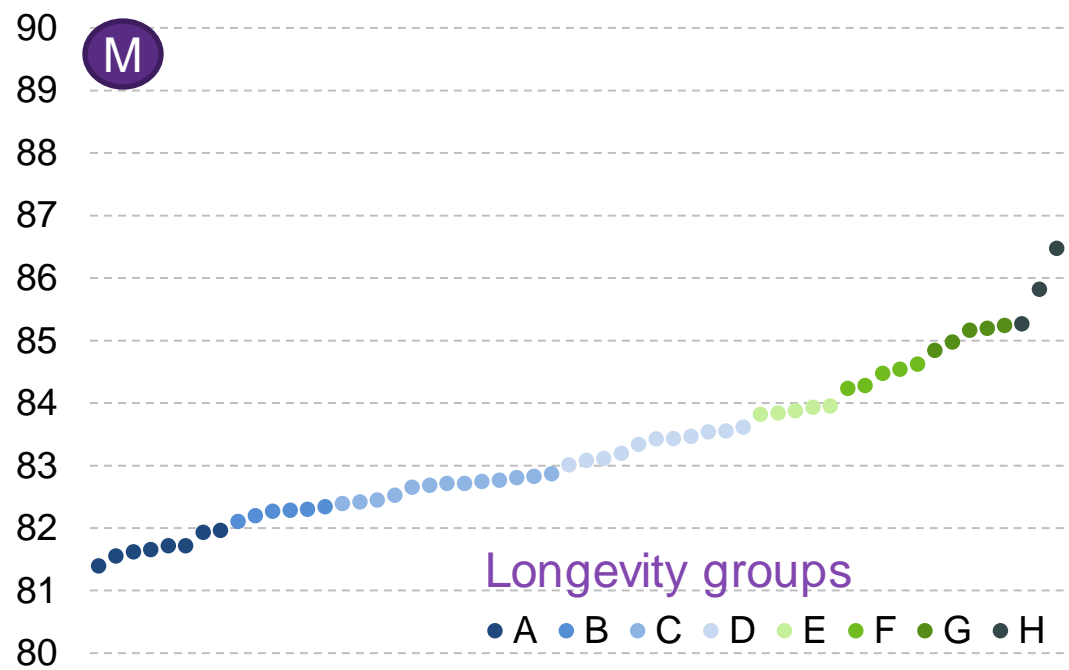
Source: Calibrating 2021 VitaCurves v2

CLUBVITA

3 How can you provide PII safely?
Brad Malin

De-identified Data Sharing: Principles and Opportunities

Bradley Malin, Ph.D.

Vanderbilt University Medical Center

Vice Chair, Department of Biomedical Informatics
Co-Director, Center for Health Data Science
Co-Director, Center for Genetic Privacy & Identity in Community Settings

Vanderbilt University

Accenture Professor of Biomedical Informatics, Biostatistics, and CS
Co-Director, Big Biomedical Data Science Training Program
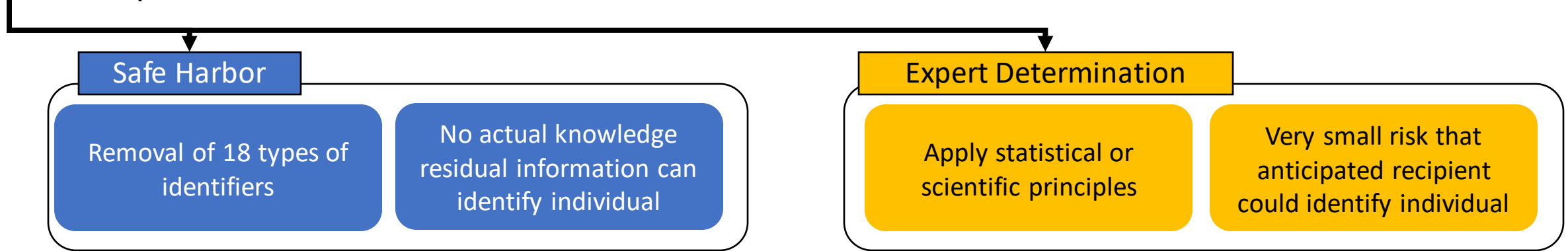
October 4, 2022

Disclaimer

- I have been a paid consultant to Club Vita.

De-identification

US Health Insurance Portability & Accountability Act (HIPAA) does not apply to de-identified data:

"information that does not identify an individual and … no reasonable basis … information can be used to identify an individual"

Safe Harbor

Removal of 18 types of identifiers	No actual knowledge residual information can identify individual

Expert Determination

Apply statistical or scientific principles	Very small risk that anticipated recipient could identify individual

There are also de-identification exemptions in the US State consumer data privacy acts

EU General Data Protection Regulation [GDPR]) does not apply to anonymous data:

"The principles of data protection should therefore not apply to anonymous information… To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."

De-identification is in Practice

- Many companies in the US and abroad apply de-identification to…
  - Support internal analytics
  - Support real world evidence investigations
  - Publish clinical trials
  - …

- Various initiatives by not-for-profits and governmental agencies do so as well

IBM MarketScan Research Databases

Take advantage of in-depth, patient-level information for easier, more accurate analyses.

Schedule a consultation

Reasons why you should use IBM MarketScan Research Database

One of the most extensive databases for research

IBM® MarketScan® Research Databases is a suite of proprietary databases that contain one of the longest-running and largest collection of privately and publicly insured, de-identified patient data in the United States. MarketScan claims data reflect the real world of treatment patterns and costs by tracking millions of patients as they travel through the healthcare system (even as they switch health plans).

The ability to capture the full continuum of care

MarketScan provides complete episodes of care support w inclusive cost and treatment studies and development of b plans. The databases cover: physician office visits, hospital order, specialty pharmacies and carve-out care, such as m

Linked data for enhanced research

MarketScan can link healthcare data from a variety of sources, including: 1) Hospital discharge records and Electronic Medical Records linked with claims data at the patient level. 2) Data on absenteeism, disability and workers' compensation – all linked to the corresponding employee's claims. 3) Linked

Detailed prescription drug information

MarketScan databases contain comprehensive information prescriptions. Through the MarketScan Early View Data Set within 45 days of the end of the service month, with month updates. MarketScan data allow identification of the type o

Safe Harbor: A Recipe for Privacy

Field	Detail
Names	Related to patient (not providers)
Unique Numbers	Phone, Social Security Number, …
Internet	Email, URL, IP addresses, …
Biometrics	Finger, voice, …
Dates	Less specific than year Ages > 89
Geocodes	Town, County, Less specific than Zip-3 (assuming > 20,000 people in zone)
Catch-All	Any other unique ID, characteristic, or code

Limited Dataset (protected under HIPAA)

Safe Harbor (Not protected under HIPAA)

A Simplified View on Risk

Controlling the Data

- Uniqueness

- Replicability

- Availability

Controlling the Setting

- Use Agreements

- Unique Login & Audits

- Pay for Access

- *Sampling*

Categorize Variables By Gross Risk

- Low risk: fail on one of the three principles of identifiability
  - E.g., Amount of co-pay incurred by an individual

- Moderate risk: may potentially identify someone
  - E.g., Age, Gender, Date of Service, Place of Service

- High risk: will certainly identify someone
  - E.g., Patient Name, Social Security Number

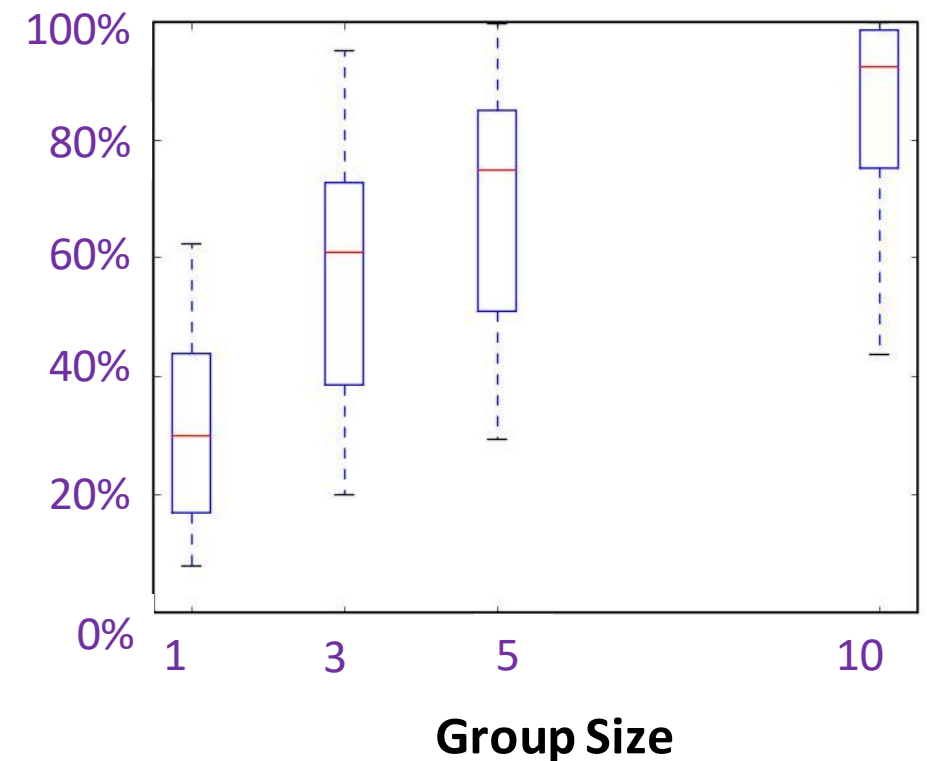- All the while – this analysis must account for capabilities of recipient

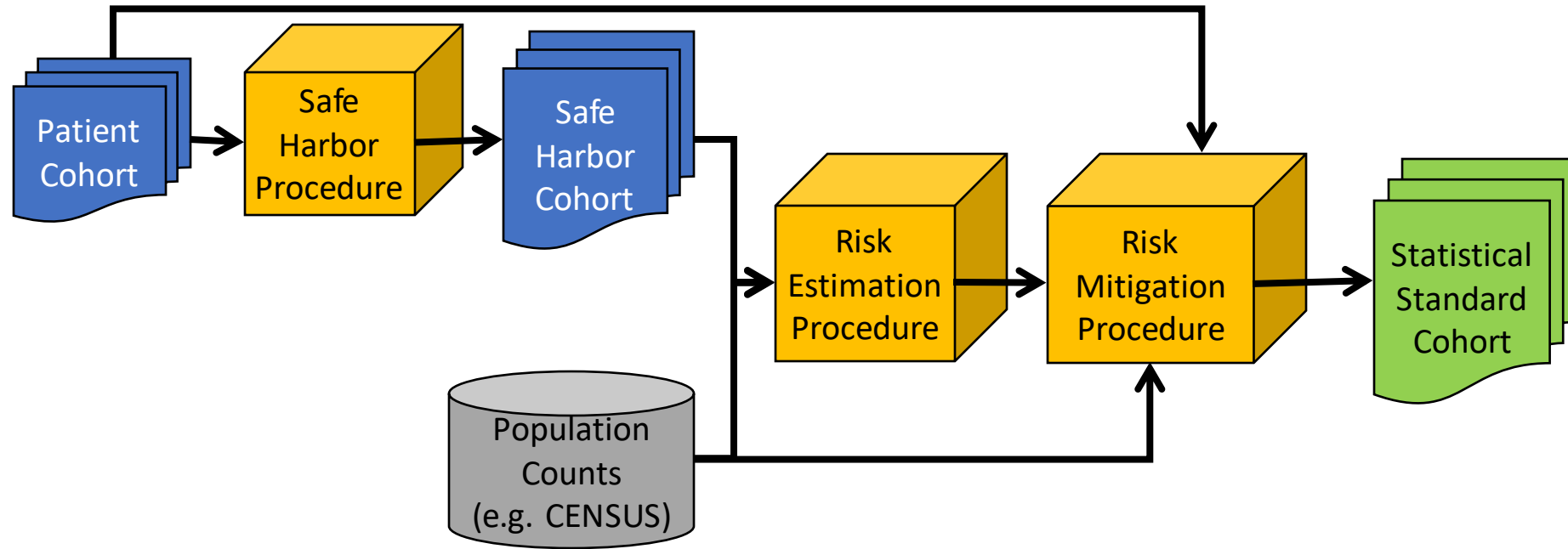# All U.S. States {Geocode, Gender, Age, Race}

Safe Harbor (De-identified)

Limited Data set
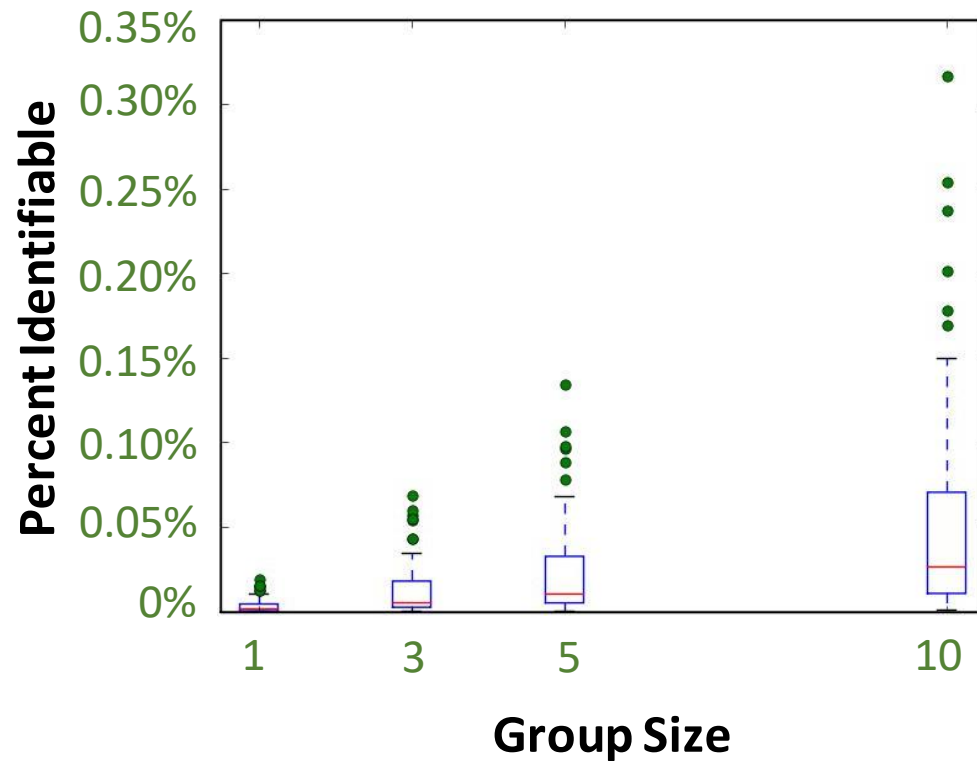
Benitez & Malin, JAMIA. 2010.

An Expert De-identification Process

Malin, Benitez, and Masys. JAMIA. 2011.
Xia, et. al. JAMIA. 2021

Vanderbilt Case Case Study

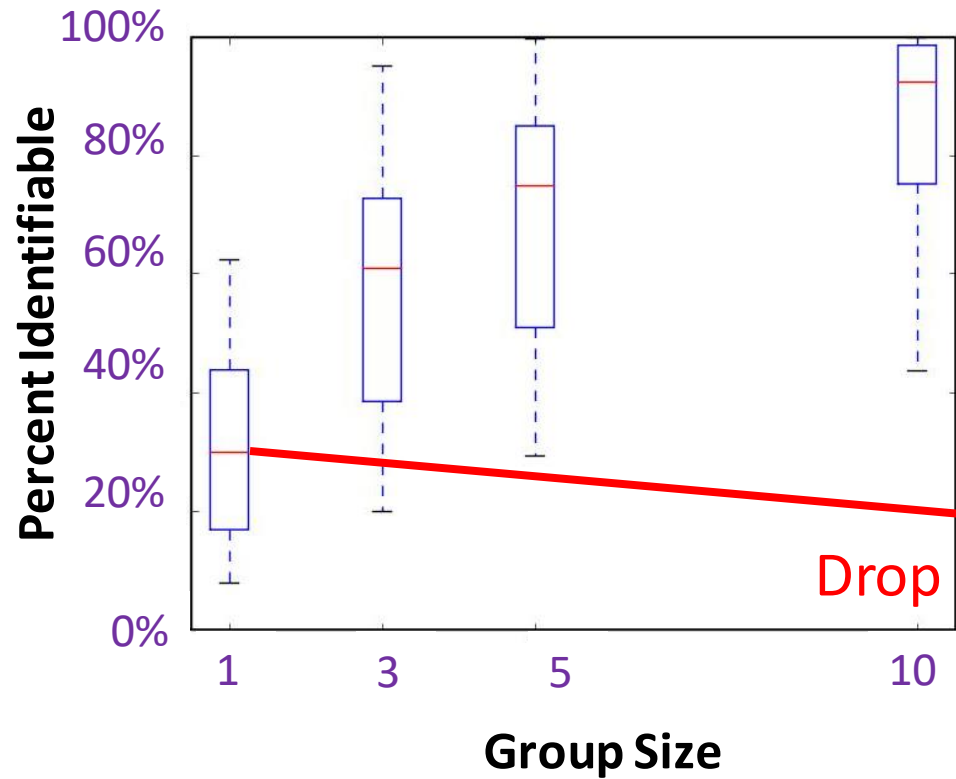| Who	State	State Population Size (US Census)	Cohort Size	Patients >89 years old
Vanderbilt	TN	~5M	~3K	12

Policy	Generalizations			Risk
	Gender	**Race**	**Age**	
Safe Harbor	∅	∅	[90 - 120]	0.909
Alternative 1	[M or F]	∅	∅	0.476
Alternative 2	∅	[Asian or Other]	∅	0.857
Alternative 3	∅	∅	[52 - 53]	0.875

The Availability of Demographics Varies...

	IL	MN	TN	WA	WI
WHO	Registered Political Committees (ANYONE – In Person)	MN Voters	Anyone	Anyone	Anyone
Format	Disk	Disk	Disk	Disk	Disk
Cost	$500	$46; "use ONLY for elections, political activities, or law enforcement"	$2500	$30	$12,500
Name	●	●	●	●	●
Address	●	●	●	●	●
Date of Birth	●	○	●	●	
Sex	●		●	●	
Race			●		
Phone Number	●	●			

Identifiability Changes!

Limited Data Set

Limited Data Set ⟵⟶ Voter Reg.

Drop by 33%!!
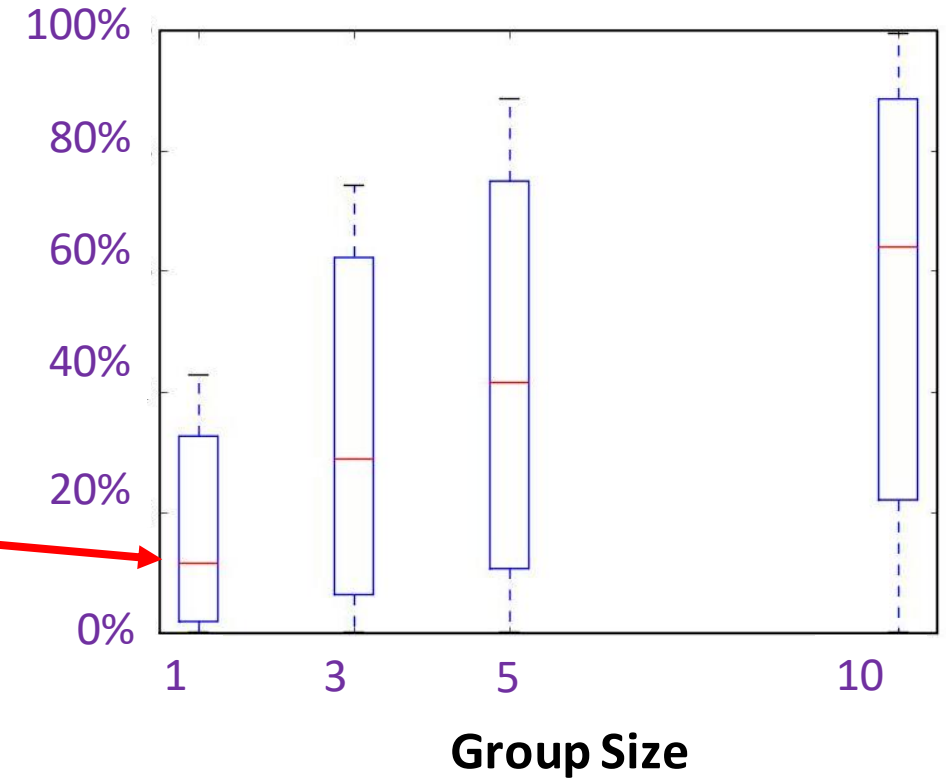
Parting Thoughts

- There is ***always*** a risk of re-identification

- But risk exists in any security setting

- You can sufficiently mitigate risk by combining data de-identification with controlled environments and contractual agreements

- There is a need for industry standards

Questions?

brad.malin@gmail.com

Center for Genetic Privacy and Identity in Community Settings
http://www.vumc.org/getprecise/

Vanderbilt Health Data Science Center
http://www.vumc.org/heads/

Vanderbilt Health Information Privacy Laboratory
http://www.hiplab.org/

Balancing privacy and best price in the age of US pension risk transfer

Danielle Prado
Panelist

Senior Associate,
MBM Commercial
LLP

Tom Jones
Panelist

Pricing Actuary,
Fidelity & Guaranty
Life Insurance
Company

Brad Malin
Panelist

Accenture Professor
of Biomedical
Informatics,
Biostatics and CD,
Vanderbilt University

Douglas Anderson
Webinar chair

Chief Visionary Officer,
Club Vita