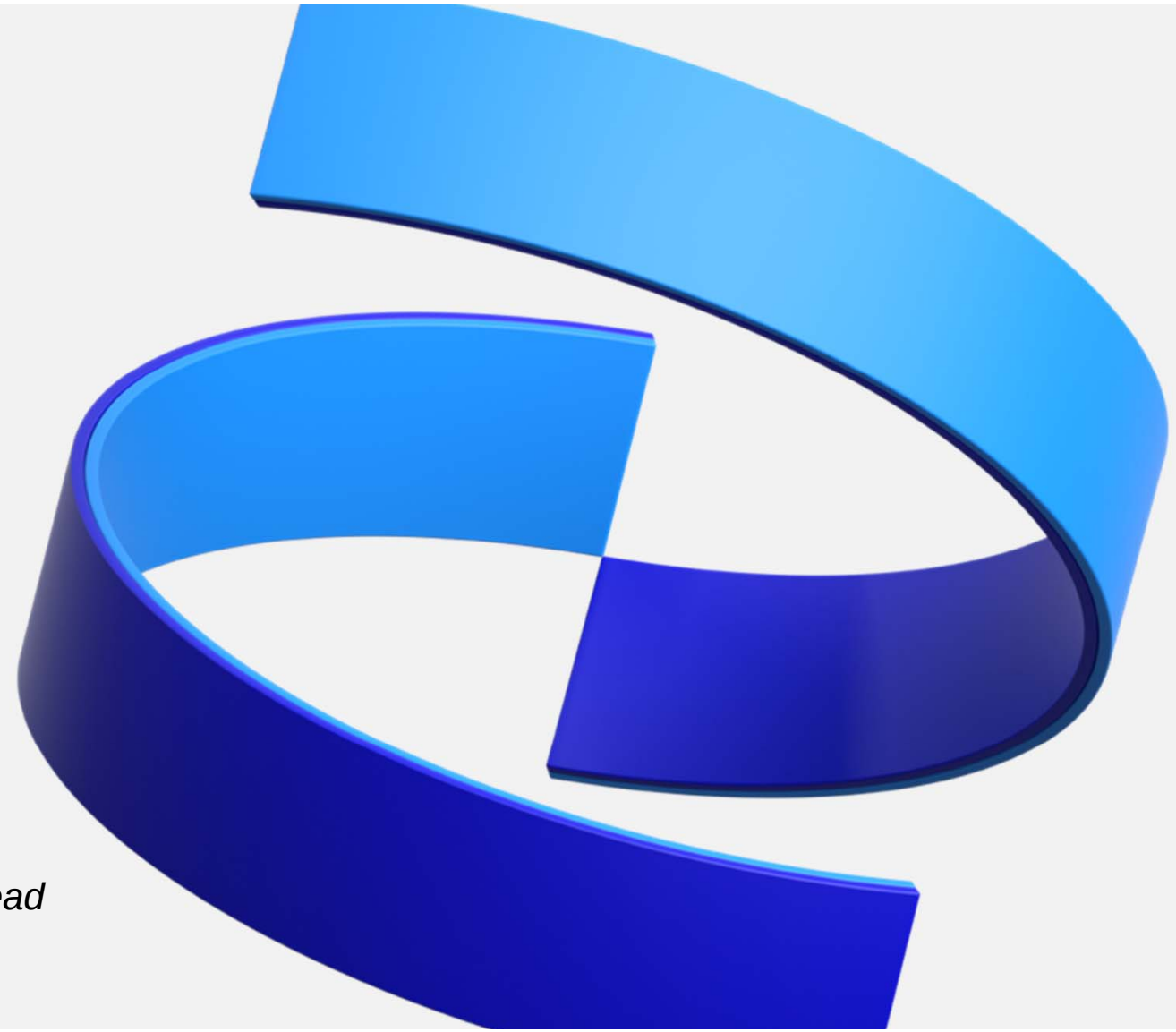
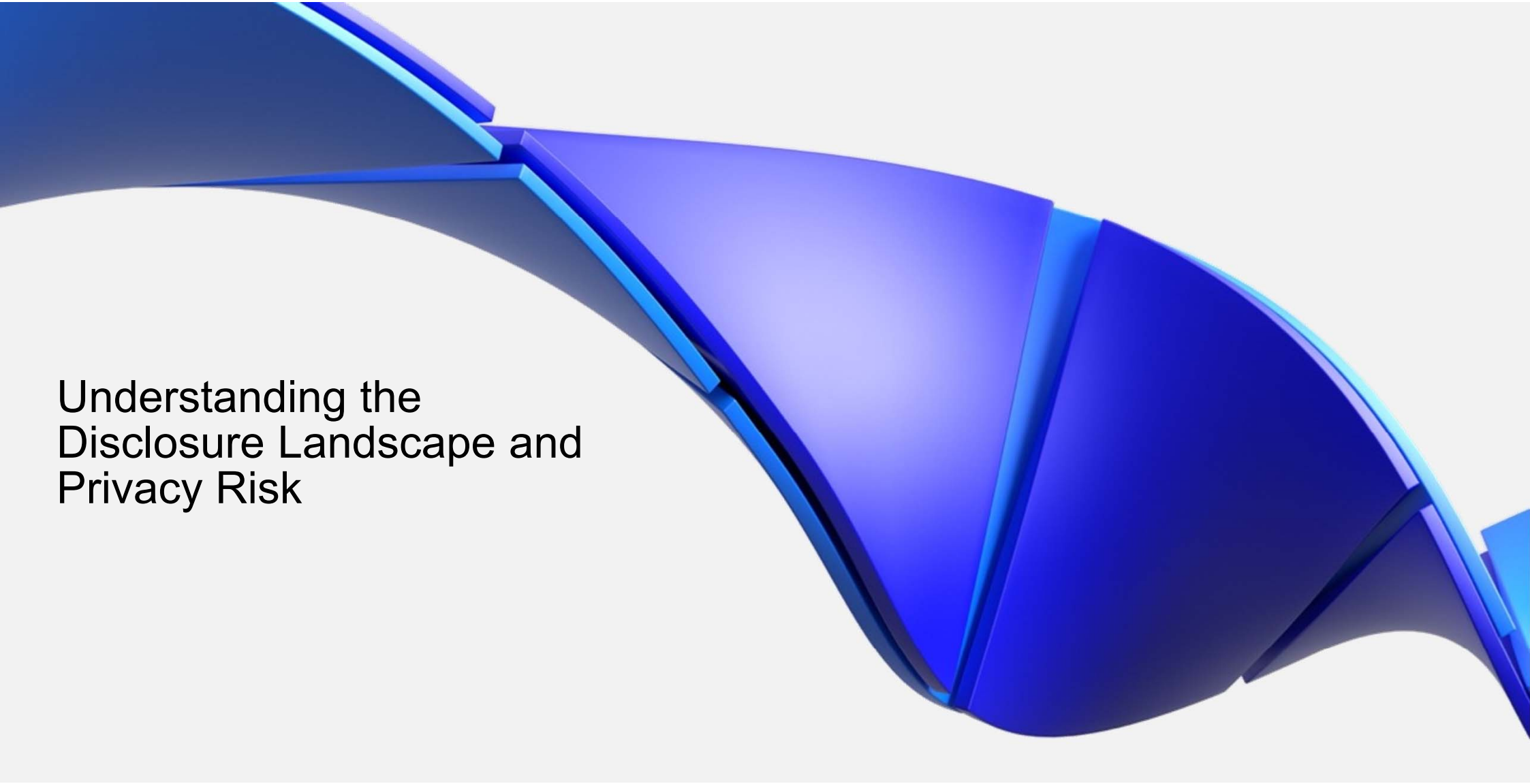
COVID Vaccine Public Disclosure



Lora Killian
Clinical Trial Transparency & Disclosure Lead
November 2021

Pfizer

Understanding the Disclosure Landscape and Privacy Risk



Assume that every document will be (or will eventually be). . .

public

The Evolving Transparency Landscape

EU Clinical Trial Regulation* (1/31/2022): requires disclosure of the clinical trial application (with a few component exceptions) and disclosure of some end of trial documents:
- Protocol
- IMPD (Sections S & E)
- IB

ClinicalTrials.gov:
- Trial Registration

ICMJE:
- Trial Registration

ClinicalTrials.gov/EudraCT:
- Basic Results
- Protocol/SAP (US only)

EU Clinical Trial Regulation* (1/31/2022):
- CSR & Technical Summary
- Plain Language Summary

ICMJE:
- Data Sharing Plan within Publications

PhRMA/EFPIA:
- CSR Synopsis
- Plain Language Summary
- Secondary Research

Regulations that can be triggered any time after documents are submitted to the regulatory agencies:
- EMA Policy 0043
- US Freedom of Information Act (upon request)
- HC PRCI Retrospective and Access to Documents

EMA Clinical Data Publication (CDP) aka Policy 0070/HC Public Release of Clinical Information (PRCI) Guideline:
- Module 2: Clinical Summaries, and Clinical Overview
- Module 5: CSRs (including Synopsis, Protocol, SAP, Sample CRF)

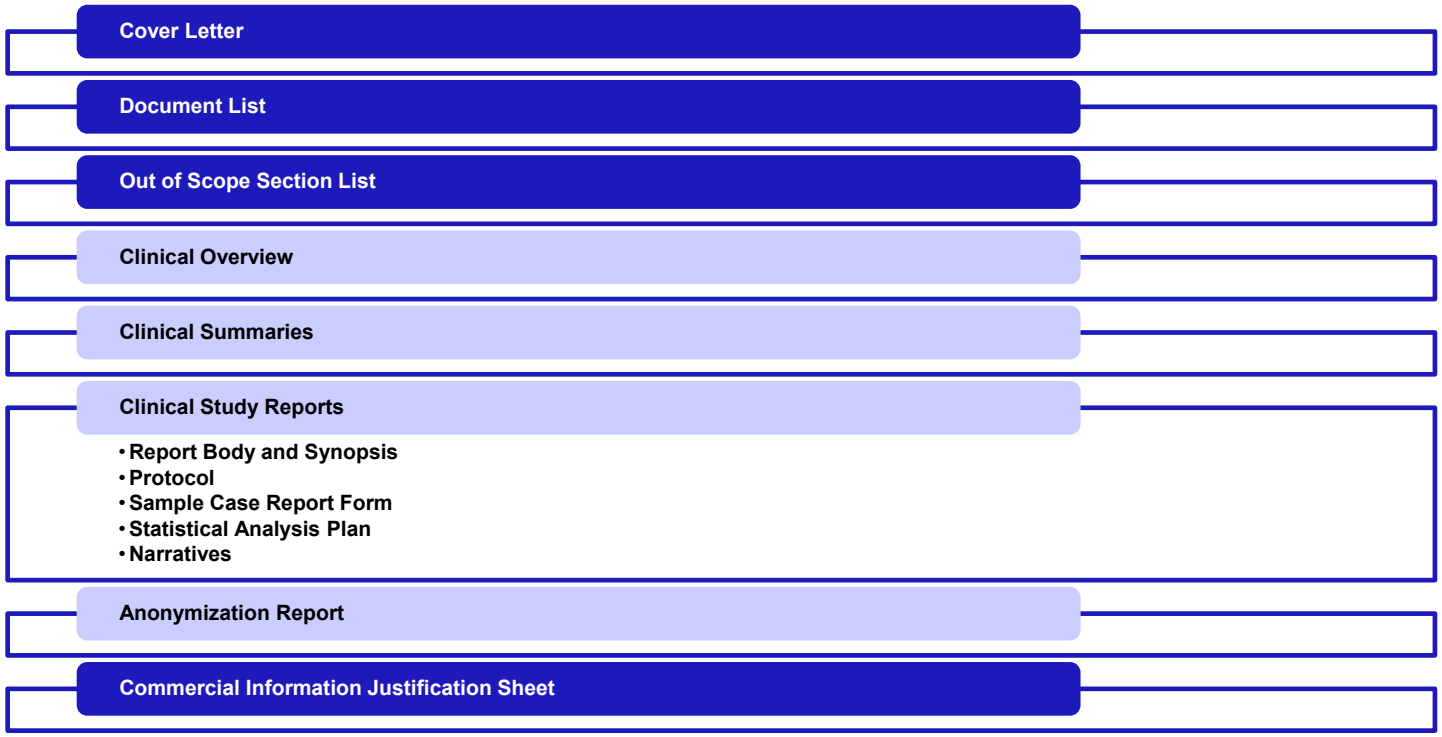
PMDA: publication policy requiring disclosure of Modules 1 and 2

*Sponsor may defer public disclosure up to the time of MA using this trial or up to 5 yr (Ph 2-3), 7 yr (Ph 1 BA, BE, biosimilar), or 1 year after the end of the trial, whichever is sooner. For Ph 4 and low-interventional trials, may defer until results posting.

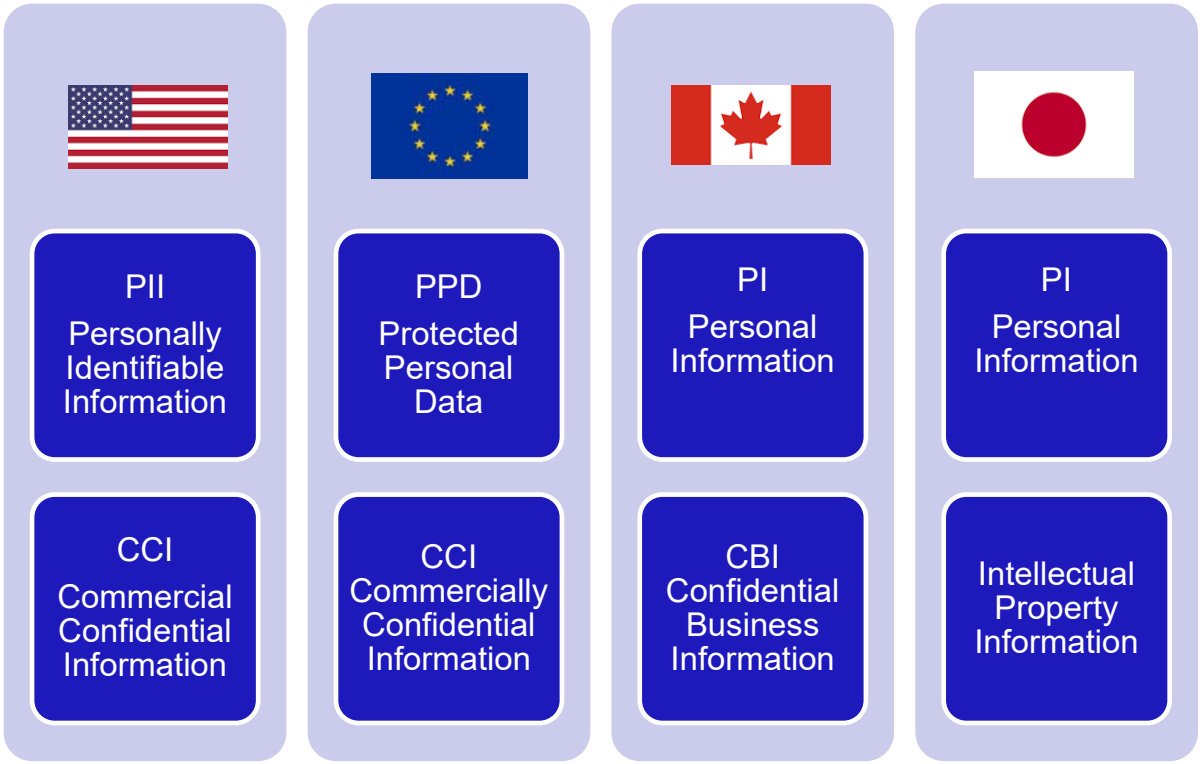
Regulatory Disclosure: HC & EMA

Documents required for submission under HC PRCI and EMA Policy 0070

Public documents must be redacted for personal and commercial information



Global Anonymization & Commercial Terms

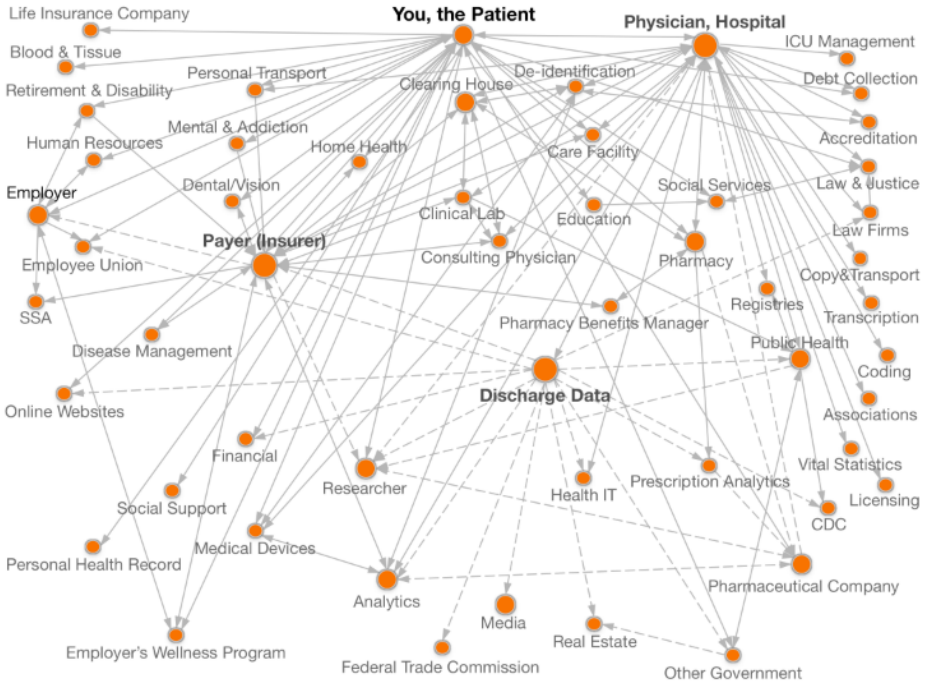


Re-identification Risk: Governor William Weld

- In 1990, Sweeney found that 87% of the US population could be identified with zip code, date of birth and gender (https://aboutmyinfo.org/)
- Mid-1990s: Massachusetts Group Insurance Commission released de-identified records to researchers for all state employees
- Governor Weld provided assurances that the information was protected because name, address, SSN and other “explicit identifiers” were removed (~100 attributes remained)
- May 18, 1996: Governor Weld collapses while delivering keynote graduation address at Bentley college
- 1997: For $20, Latanya Sweeney purchased voter registration records; used the records to identify Governor Weld in the publicly available data

Medical Information in the Public Domain

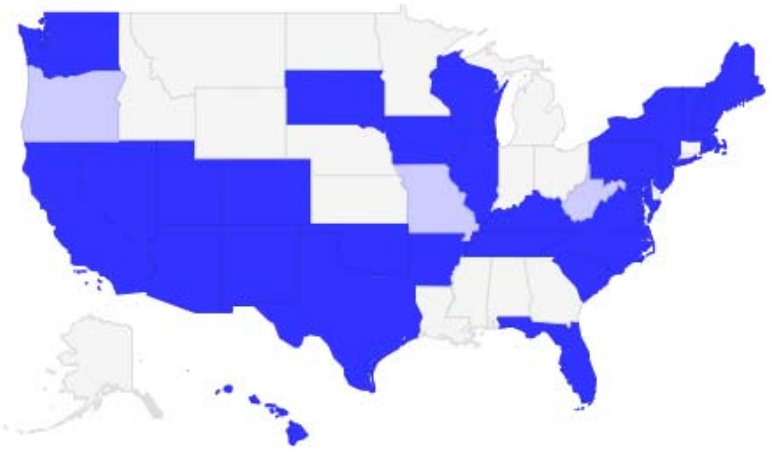
Tracking Medical Information



https://thedatamap.org/map2013/p28.php

Sharing of Medical Data in the US

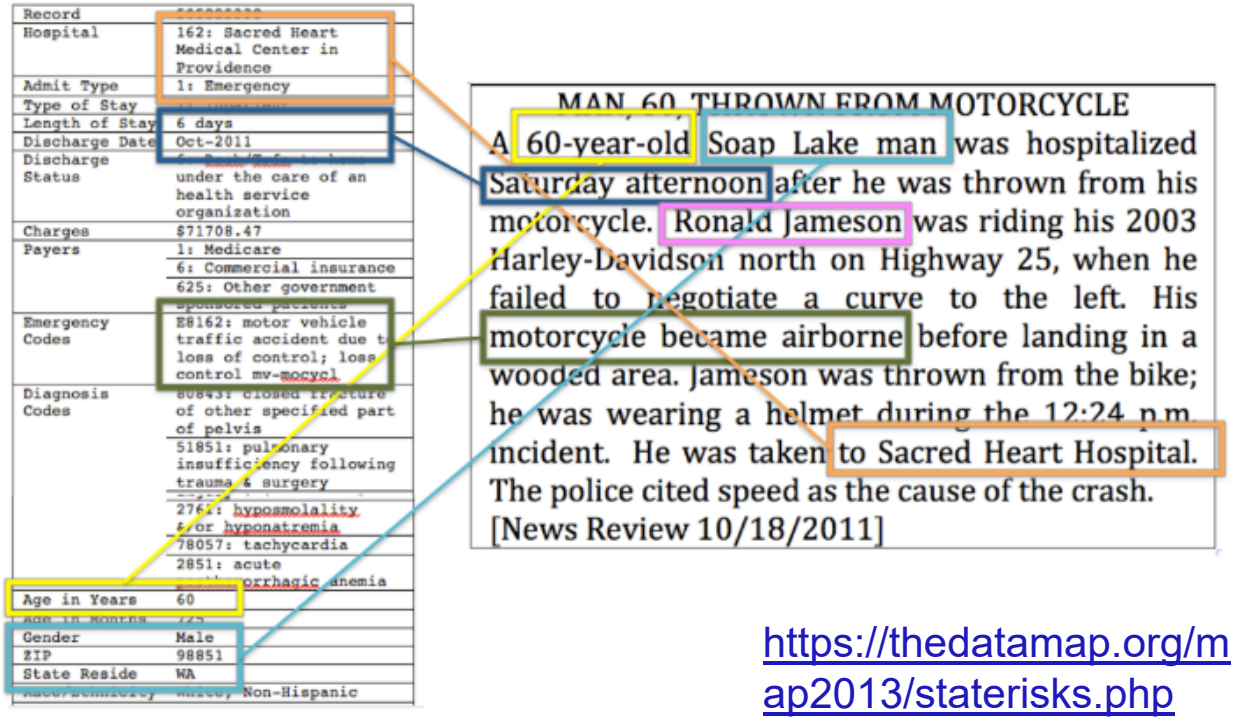
- 33 states sell or give away personal health data (all blue shaded states)
- Only 3 states use protections as strict as HIPAA (light blue states)



http://thedatamap.org/map2013/states.php

What information creates risk?

- Location (investigator sites are increasingly required to be made public)
- Medical History
- Concomitant Medications
- Adverse Events
- Lab Values
- Diagnostic Values



Record	
Hospital	162: Sacred Heart Medical Center in Providence
Admit Type	1: Emergency
Type of Stay	
Length of Stay	6 days
Discharge Date	Oct-2011
Discharge Status	under the care of an health service organization
Charges	$71708.47
Payers	1: Medicare 6: Commercial insurance 625: Other government
Emergency Codes	E8162: motor vehicle traffic accident due t loss of control; loss control mv-mocycl
Diagnosis Codes	80843: closed fracture of other specified part of pelvis 51851: pulmonary insufficiency following trauma & surgery 2761: hyposmolality &/or hyponatremia 78057: tachycardia 2851: acute [illegible]orrhagic anemia
Age in Years	60
Age in Months	725
Gender	Male
ZIP	98851
State Reside	WA
	Non-Hispanic

MAN, 60, THROWN FROM MOTORCYCLE

A 60-year-old Soap Lake man was hospitalized Saturday afternoon after he was thrown from his motorcycle. Ronald Jameson was riding his 2003 Harley-Davidson north on Highway 25, when he failed to negotiate a curve to the left. His motorcycle became airborne before landing in a wooded area. Jameson was thrown from the bike; he was wearing a helmet during the 12:24 p.m. incident. He was taken to Sacred Heart Hospital. The police cited speed as the cause of the crash. [News Review 10/18/2011]

https://thedatamap.org/map2013/staterisks.php

Example Patient Narratives

Yellow highlighted text indicates potential privacy issues

Marked for Redaction Version

This is a report for Protocol XYZ1234, Center ID|Subject ID 1000.

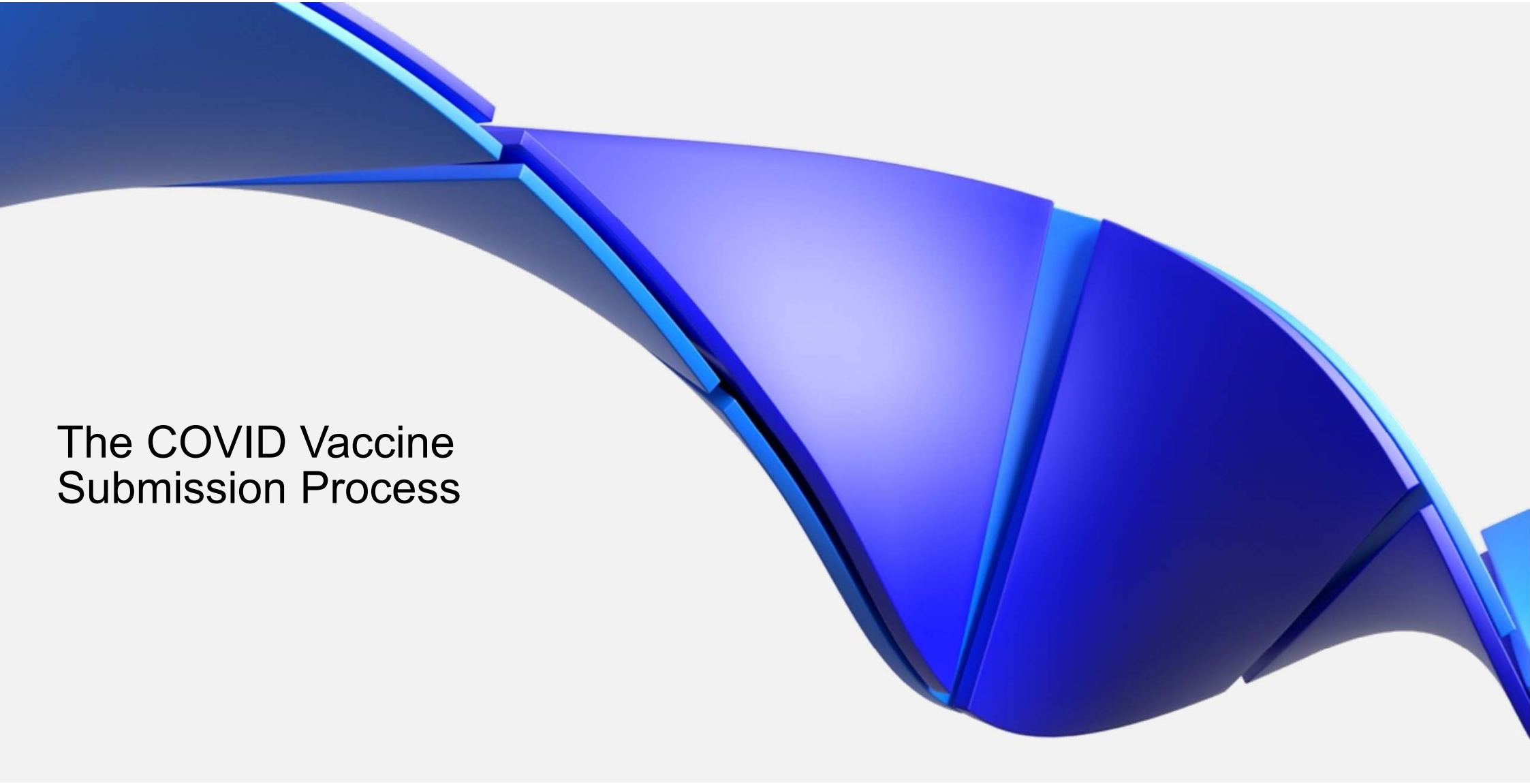
A 50-year-old female subject (unknown ethnicity) was enrolled on 20April2010 in the above mentioned study. Medical history included hypothyroidism, hypercholesterolemia, appendicectomy, partial mastectomy in 1993 and single cardiac pacemaker insertion on 30May2010. As part of the study protocol, the subject received the induction phase of the study protocol as follows (for disease indication): DRUGABC from 20April2010 to 26April2010 (day 1, day 4, day 7), ribociclib (KISQALI), 600 mg/d, po, from 20April2010 to 22April2010 (day 1 through day 3) and cytarabine (ARACYTINE), 200 mg/m2/d, IV, from 20April2010 to 26April2010 (day 1 through day 7). After the induction phase, the subject developed symptomatic bradycardia with sinus dysfunction, leading to right ventricular pacemaker insertion on 30May2010. The subject was discharged on 07June2010 and then received DRUGABC. On 14July2010, the subject was admitted to the Hematology Unit for the first consolidation phase which took place on 16July2010 after catheter placement on the previous day. As part of the first consolidation treatment, the subject received DRUGABC, once a day, on 16July2010 (Day 1), ribociclib, 600 mg/ day, po from 16July2010 to 17July2010 (Day 1) and cytarabine (ARACYTINE), 1 g per day, by intravenous drip route from 16July2010 to 19July2010 (Day 1 to Day 4). Concomitant treatments included levothyroxine sodium (SYNTHROID), hydroxyzine hydrochloride (ATARAX), valaciclovir (ZELITREX) and venlafaxine (Effexor XR). The subject tolerated well the chemotherapy with some nausea, cramps and mucositis. Toxidermia of the trunk due to ARACYTINE was diagnosed on 23July2010. On 26July2010, the subject presented fever 38.1°C, chills and marbled skin. Antibiotic treatment of cefepim (AXEPIM) was initiated (dose not specified). On 27July2010, the subject experienced septic shock due to Streptococcus and head trauma after a fall while stepping out of the shower without loss of consciousness (both life-threatening) and with cranial trauma. The septic shock led to hemodynamic disorder with hypotension, which may be related with the subject fall. The septicemia occurred while the subject was in bone marrow aplasia. Its treatment included the imipenem/cilastatine (TIENAM), vancomycine, gentamycine, amphotericine B (AMBISOME). On 27July2010 platelet count was 25000/mm3 (N: 150000 - 400000), prothrombin level was 61 % (N: 70-100), GGT was 270 UI/L (N: > 35), white blood cell count was 400/mm3 (N: 4000 - 10000). On 29July2010 platelet count was 11000 /mm3 (N: 150000 / 400000), blood potassium was 2.9 mol/l (N: 3.5 / 5). On 26July2010 blood culture showed Cocci Gram + (Streptococcus). The subject had recovered from septic shock due to streptococcus on 29July2010. The subject had recovered from head trauma after fall while stepping out of the shower without loss of consciousness on an unspecified date. The subject was discharged on 29July2010 after 48 hours in the Critical Care Unit.

Redacted Version

This is a report for Protocol XYZ1234, Center ID|Subject ID [redacted].

A [redacted]-year-old [redacted] subject (unknown ethnicity) was enrolled on [redacted] in the above mentioned study. Medical history included hypothyroidism, hypercholesterolemia, appendicectomy, partial mastectomy in [redacted] and single cardiac pacemaker insertion on [redacted]. As part of the study protocol, the subject received the induction phase of the study protocol as follows (for disease indication): DRUGABC from [redacted] to [redacted] (day 1, day 4, day 7), ribociclib (KISQALI), 600 mg/d, po, from [redacted] to [redacted] (day 1 through day 3) and cytarabine (ARACYTINE), 200 mg/m2/d, IV, from [redacted] to [redacted] (day 1 through day 7). After the induction phase, the subject developed symptomatic bradycardia with sinus dysfunction, leading to right ventricular pacemaker insertion on [redacted]. The subject was discharged on [redacted] and then received DRUGABC. On [redacted], the subject was admitted to the Hematology Unit for the first consolidation phase which took place on 16July2010 after catheter placement on the previous day. As part of the first consolidation treatment, the subject received DRUGABC, once a day, on [redacted] (Day 1), ribociclib, 600 mg/ day, po from [redacted] to [redacted] (Day 1) and cytarabine (ARACYTINE), 1 g per day, by intravenous drip route from [redacted] to [redacted] (Day 1 to Day 4). Concomitant treatments included levothyroxine sodium (SYNTHROID), hydroxyzine hydrochloride (ATARAX), valaciclovir (ZELITREX) and venlafaxine (Effexor XR). The subject tolerated well the chemotherapy with some nausea, cramps and mucositis. Toxidermia of the trunk due to ARACYTINE was diagnosed on [redacted]. On [redacted], the subject presented fever 38.1°C, chills and marbled skin. Antibiotic treatment of cefepim (AXEPIM) was initiated (dose not specified). On [redacted], the subject experienced septic shock due to Streptococcus and head trauma after a fall while stepping out of the shower without loss of consciousness (both life-threatening) and with cranial trauma. The septic shock led to hemodynamic disorder with hypotension, which may be related with the subject fall. The septicemia occurred while the subject was in bone marrow aplasia. Its treatment included the imipenem/cilastatine (TIENAM), vancomycine, gentamycine, amphotericine B (AMBISOME). On [redacted] platelet count was 25000/mm3 (N: 150000 - 400000), prothrombin level was 61 % (N: 70-100), GGT was 270 UI/L (N: > 35), white blood cell count was 400/mm3 (N: 4000 - 10000). On [redacted] platelet count was 11000 /mm3 (N: 150000 / 400000), blood potassium was 2.9 mol/l (N: 3.5 / 5). On [redacted] blood culture showed Cocci Gram + (Streptococcus). The subject had recovered from septic shock due to streptococcus on [redacted]. The subject had recovered from head trauma after fall while stepping out of the shower without loss of consciousness on an unspecified date. The subject was discharged on [redacted] after 48 hours in the Critical Care Unit.

The COVID Vaccine Submission Process



Setting the Scene:

- Timelines:
  - HC EUA Submitted – 10/9/2020
  - EMA MAA Submitted – 11/30/2020
    (Rolling review started 10/6/2020)
  - HC EUA Decision – 12/9/2020
  - EMA MAA Decision – 12/21/2020
- Context:
  - Pandemic
  - BioNTech/Pfizer Partnership
  - Pfizer's traditional approach to anonymization
  - Holidays fast approaching ☺

Prior Pfizer PRCI Submissions to Health Canada

- Pfizer's current approach to HC's Public Release of Clinical Information (PRCI) anonymization includes full redaction of participant narratives
- 3 Submissions delivered in 2020 with fully redacted narratives (dates below reflect publication date):
  - Mylotarg (8/14/2020)
  - Vyndaqel (12/11/2020)
  - Daurismo (3/1/2021)
- All three publication packages received non-conformance statements

Duarismo Health Canada Statement:

NOTICE:

This clinical information package includes extensive redactions to the patient information and/or data listed below. These specific redactions do not conform to Health Canada guidance, which encourages manufacturers to retain the analytical value of information by using other transformation methods (e.g., generalization or randomization), and to apply these methods to specific information that risks re-identifying an individual rather than to redact broad sections of information.

List of redacted information:
Entire documents

- Tables 14.3.3.1 Death Narratives
- Tables 14.3.3.2 Other Serious adverse event narratives
- Tables 14.3.3.3.1 Non-serious adverse events
- Tables 14.3.3.3.2 Adverse events of special interest

Within report bodies, portions of redacted information pertaining to the same categories of information as listed above.

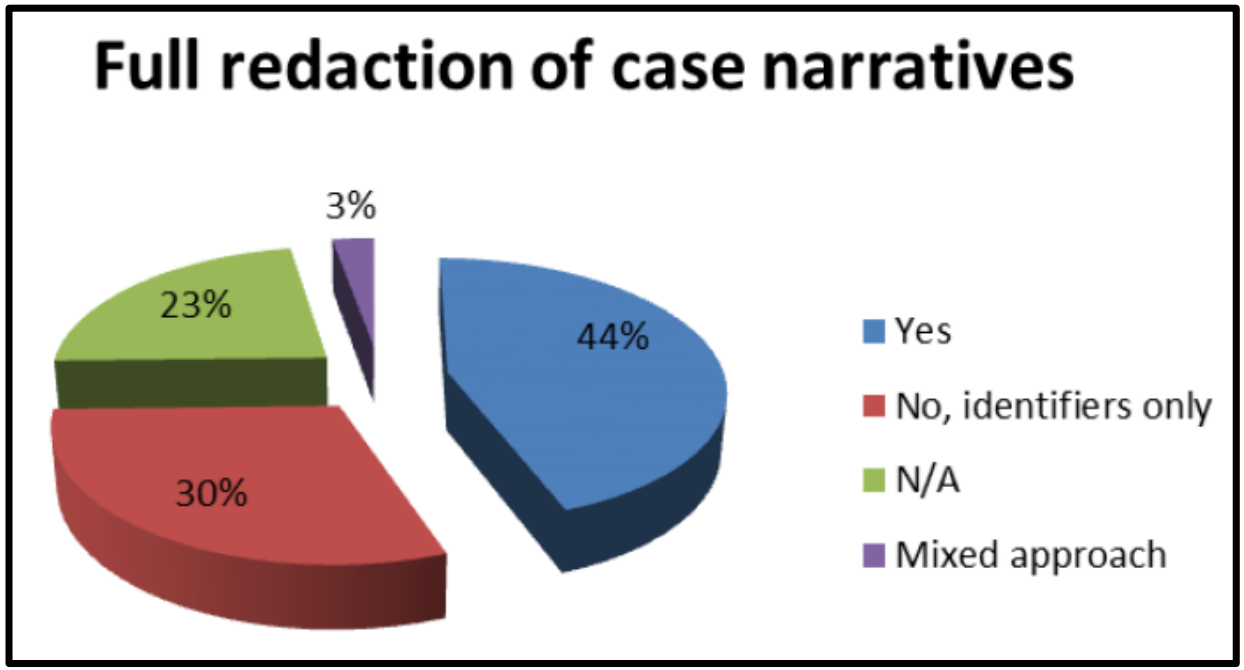
Health Canada encourages manufacturers to anonymize personal information according to the principles outlined in Guidance Document: Public Release of Clinical Information. Health Canada will continue to explore ways to help ensure all publications include anonymized clinical information.

If you require access to the redacted information, you may submit inquiries to the Information Science and Openness Division (hc.clinicaldata-donneescliniques.sc@canada.ca).

EMA Clinical Data Publication (CDP) Policy

EMA presentation from December 2019

- 142 published procedures
- 7089 document published
- Nearly 4.4 million pages published
- Policy paused in October 2018



COVID Disclosure Planning: EMA CDP and HC PRCI

European Medicines Agency (EMA)
Clinical Data Publication (CDP)

- Original Publication Timeline for COVID Products: ASAP (EMA)
- Pfizer/BioNTech/EMA 10/16/2020 Meeting:
  - No specific date given but historically process did not begin until a year post EC decision
  - EMA would like to move forward with joint calls and review with HC
  - EPAR initial publication proposed for 3 days post decision
  - Full RMP will be published ASAP once decision rendered
  - EMA has accepted block (100%) redaction of narratives in previous submissions from Pfizer

Health Canada (HC)
Public Release of Clinical Information (PRCI)

- Original Publication Timeline for COVID Products: 120 Days from interim order authorization (HC)
- Pfizer/Health Canada 12/3/2020 Meeting:
  - HC would like to publicly post the submission by February 1st
  - HC would like to move forward with joint calls and review with the EMA
  - HC stated that they do not want block (100%) redaction of narratives; HC considers a participant's medical information to have low risk of re-identification and as result is not personal identifiable information

COVID-19 Vaccine Submission Document List

24 in-scope documents for a total of 12,210 pages

File Name	In-Scope EMA	In-Scope HC	Page Count
Total			**12,210**
Study BNT162-01			**4,132**
bnt162-01-interim-notes-for-reader	Yes	Yes	3
bnt162-01-interim-protocol	No	Yes	292
bnt162-01-interim-report-body	No	Yes	1,115
bnt162-01-interim-sample-crf	Yes	Yes	139
bnt162-01-interim-sap	No	Yes	33
bnt162-01-interim-synopsis	No	Yes	14
bnt162-01-interim2-protocol	Yes	No	338
bnt162-01-interim2-report-body	Yes	No	2,145
bnt162-01-interim2-sap	Yes	No	35
bnt162-01-interim2-synopsis	Yes	No	18
Study C4591001			**7,414**
c4591001-fa-interim-errata	Yes	Yes	1
c4591001-fa-interim-protocol	Yes	Yes	1,413
c4591001-fa-interim-report-body	Yes	Yes	2,033
c4591001-fa-interim-sample-crf	Yes	Yes	212
c4591001-fa-interim-sap	Yes	Yes	59
c4591001-fa-interim-synopsis	Yes	Yes	31
c4591001-Final Analysis Interim Narrative (Sensitive)	Yes – Delayed	Yes – Delayed	3,611
Safety Narratives Subset Unblinded	No	Yes – Delayed	16
Severe COVID-10 Efficacy Narratives Subset Unblinded	No	Yes – Delayed	38
Module 2 Documents			**664**
COVID-19 Vaccine 2.5 Clinical Overview - Initial Adult MAA Submission	Yes	Yes	257
Dec 2020 COVID-19 Vaccine MAA 2.5 Clinical Overview Appendix- Justification for Module 5 Components Not Submitted	Yes	Yes	3
EMA 2020 COVID-19-PFE and BNT assays-2.7.1 Summary of Biopharmaceutics and Associated Analytical Methods	Yes	Yes	3
COVID-19 Vaccine 2.7.3 SCE Initial Adult MAA Submission (2020)	Yes	Yes	174
COVID-19 Vaccine 2.7.4 SCS MAA Submission (2020)	Yes	Yes	227

COVID-19 Vaccine Disclosure Deliverables

Disclosure requirements to date for the COVID Vaccine Submission:

Primary Disclosure Deliverables

- Health Canada (HC) Public Release of Clinical Information (PRCI)
  - Modules 2 and 5
  - Authorization: 12/9/2020
  - Requested delivery late January for February publication
  - Delivered beginning of March
- EMA Clinical Data Publication (CDP) / (aka EMA Policy 0070)
  - Modules 2 and 5
  - Authorization: 12/21/20
  - Requested delivery late January for February publication
  - Delivered beginning of March
- Japan Disclosure
  - Modules 1 and 2
  - Authorization: 2/14/2021
  - Requested delivery mid-February
  - Delivered end of March

Secondary Disclosure Deliverables

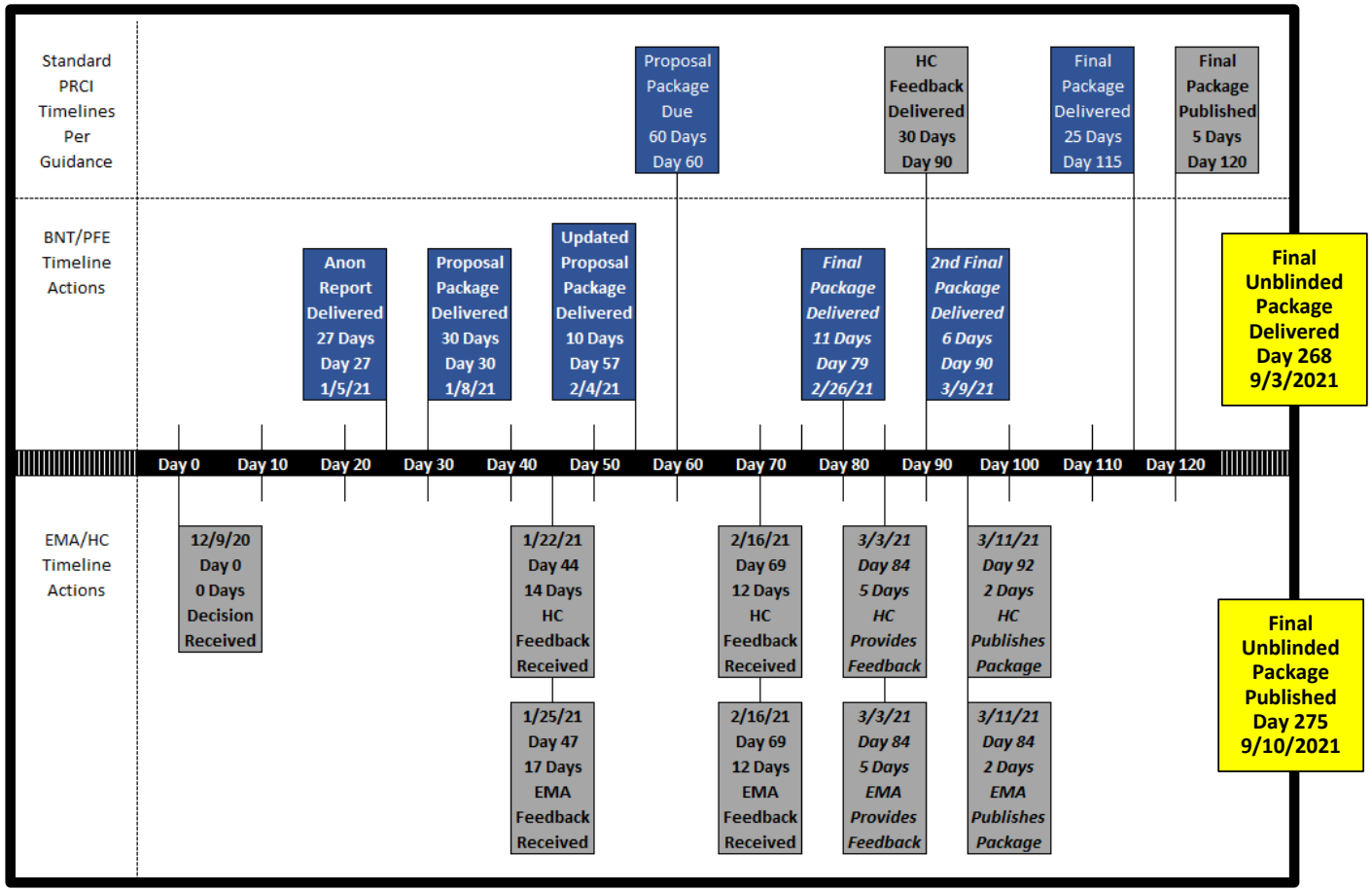
- 12/06/20: UK Assessment Report
- 12/07/20: FDA VRBPAC Briefing Document
- 12/21/20: EU RMP
- 12/21/20: EU Assessment Report
- 12/31/20: World Health Organization Approval Letter
- 01/18/21: EMA Policy 0043 - #1
- 01/21/21: Swiss Medic Freedom of Information Act
- 02/02/21: Japan Assessment Report
- 02/04/21: Canada RMP Addendum
- 02/12/21: EMA Policy 0043 - #2
- 02/22/21: EMA Policy 0043 - #3

Pfizer

Delivery Timelines

Timelines significantly shortened and a total of 6 rounds of delivery (standard is 2)

- Proposal Package delivered at **Day 30**
  (Standard: Day 60)
- EMA requested BNT 2nd Interim Analysis CSR on Day 37
- Two rounds of agency review conducted
  (Standard: 1 round of review)
  - Anonymization Report:
    - 1st Round = 74 comments
      (EMA = 50 / HC = 24)
    - 2nd Round = 52 comments
      (EMA = 29 / HC = 23)
  - Commercially Confidential Information: two rounds of review that were not coordinated with one additional informal clarification round.
- Final Package delivered at Day 79 / 11 days after review
  (Standard: Day 115 / 25 days after review)
- 2nd Final Package delivered at Day 90 / 6 days after feedback received



Blinding Concerns

Applying redactions to maintain blinding was a unique requirement

Specific Challenges with Blinding

- Created one additional full review and several updates
- Trial data characteristics posed significant unblinding risk:
  - 65% of adverse events were experienced by participants in only one treatment group
  - 47% of medical history diagnoses were experienced by participants in only one treatment group
- Cross comparison against publicly available sources difficult
- Sub-group analysis enabled unblinding when cross-referenced against the total summary analysis
- Agencies did not agree with our broader approach to fully redact complete AE, Medical History and other tables
  - BioNTech/Pfizer Team felt that the risk for error and difference in utility was not a balanced trade off for a more precise approach
  - Additionally, blinding redactions temporary - we committed to providing an updated submission package with blinding redactions removed

Public Sources of Information

- BioNTech/Pfizer VRBPAC Briefing Document (12/10/2020)
- FDA VRBPAC Briefing Document (12/10/2021)
- CHMP Assessment Report (2/19/2021)
- Publication: Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine (12/31/2020)

Tables Types Containing Blinding Redactions

- Adverse Events
- Medical History
- Concomitant Vaccines
- Demographic
- Baseline Charlson Comorbidities
- Disposition

Pfizer

Challenges with COVID Participant Narratives

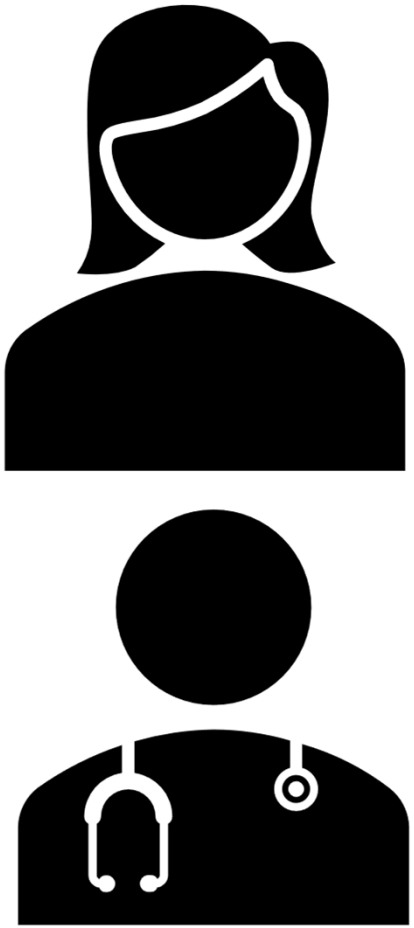
COVID-19 Vaccine Trial in the News

Source: Information and Photo from Local Newspaper

- BioNTech/Pfizer trial participant*: Dorothy Alexander
  - 41-year-old Female
  - ¼ Cherokee – Native American (stated in article)
  - 1st injection August / 2nd injection September
  - After second dose experienced temperature of 100.7 degrees
  - “My experience has been positive, and I haven't spoken to another participant who has had any issues. I am high risk for COVID, having asthma and other respiratory issues, and I have had no complications.”

Source: Information and Photo from National Newspaper

- BioNTech/Pfizer trial participant*: Allison Jones
  - 57-year-old Female
  - Black (stated in article)
  - Physician
  - Birmingham, Alabama
- BioNTech/Pfizer trial participant: Tim Smith
  - 49-year-old Male
  - State Senator from New Jersey



*Participant information anonymized to protect participant privacy

COVID-19 Vaccine Trial in the News

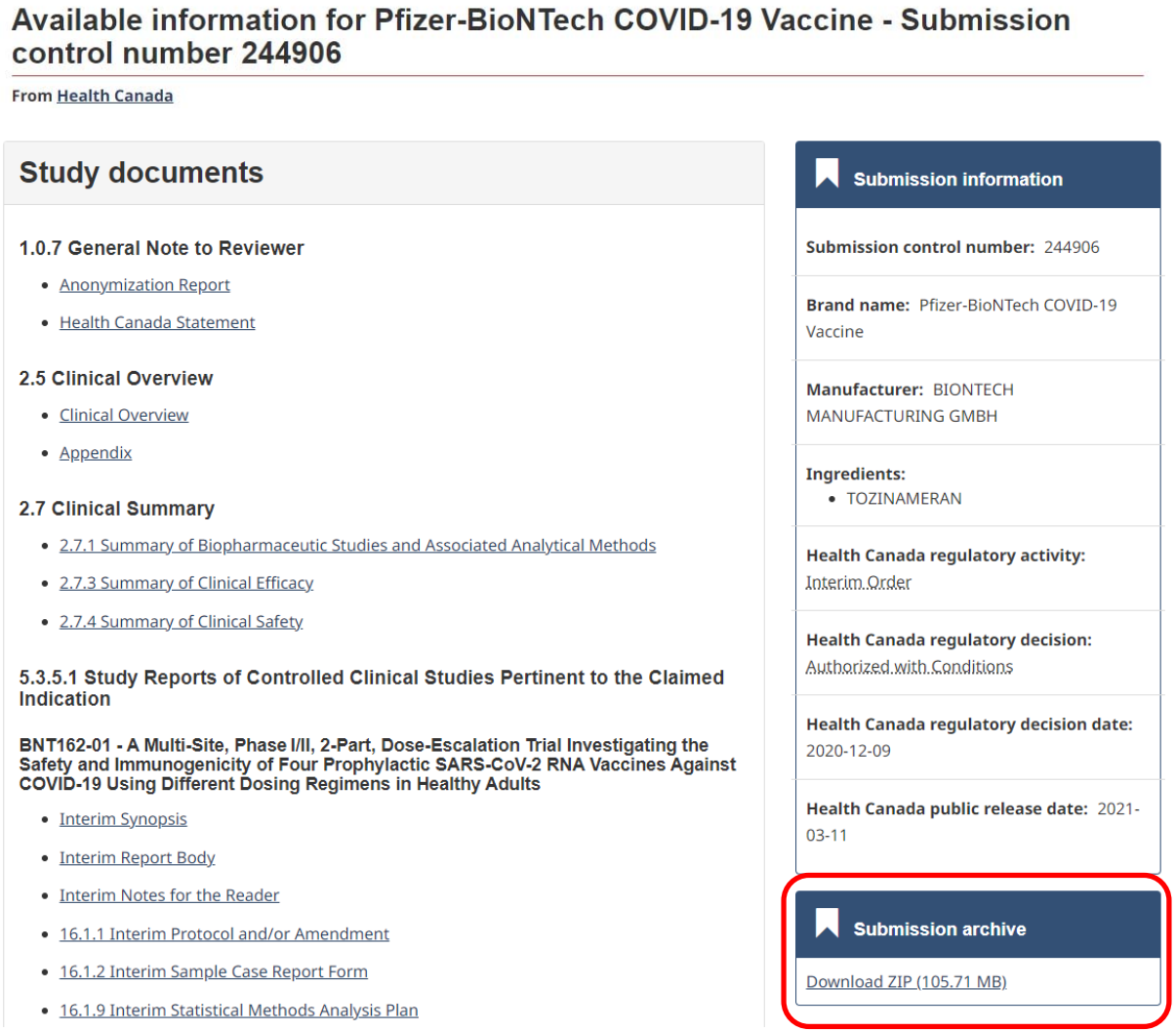
Source: Photo and Information from Personal Essay in Online Publication

- BioNTech/Pfizer trial participant: Allison Johnson
  - 53-year-old Female
  - Former fire fighter
  - “There are three facilities in the Atlanta area that are doing the study. I'm part of the one by the Clinical Research Atlanta in Stockbridge, Georgia.”
  - “I wanted to help, even though I'm high-risk for COVID-19. I have asthma and *RARE AUTOIMMUNE DISEASE.*”
  - “The next day I was exhausted, and then I developed a fever. My temperature was 100 and went up to 102 degrees for three days.”

COVID-19 Public Disclosure Submission Access

Reminder of platform differences (HC vs EMA)

- Canada site: with one click, the full submission can be downloaded from the Canadian site (see red box in image to the right)
- European site has more restrictions:
  - Requires an account to be create
  - Can only view the submission with basic account
  - Only EU citizens and those with EU affiliation can create more advanced account requiring more information to enable downloading



Available information for Pfizer-BioNTech COVID-19 Vaccine - Submission control number 244906

From Health Canada

Study documents

1.0.7 General Note to Reviewer

- Anonymization Report
- Health Canada Statement

2.5 Clinical Overview

- Clinical Overview
- Appendix

2.7 Clinical Summary

- 2.7.1 Summary of Biopharmaceutic Studies and Associated Analytical Methods
- 2.7.3 Summary of Clinical Efficacy
- 2.7.4 Summary of Clinical Safety

5.3.5.1 Study Reports of Controlled Clinical Studies Pertinent to the Claimed Indication

BNT162-01 - A Multi-Site, Phase I/II, 2-Part, Dose-Escalation Trial Investigating the Safety and Immunogenicity of Four Prophylactic SARS-CoV-2 RNA Vaccines Against COVID-19 Using Different Dosing Regimens in Healthy Adults

- Interim Synopsis
- Interim Report Body
- Interim Notes for the Reader
- 16.1.1 Interim Protocol and/or Amendment
- 16.1.2 Interim Sample Case Report Form
- 16.1.9 Interim Statistical Methods Analysis Plan

Submission information

Submission control number: 244906

Brand name: Pfizer-BioNTech COVID-19 Vaccine

Manufacturer: BIONTECH MANUFACTURING GMBH

Ingredients:
- TOZINAMERAN

Health Canada regulatory activity: Interim Order

Health Canada regulatory decision: Authorized with Conditions

Health Canada regulatory decision date: 2020-12-09

Health Canada public release date: 2021-03-11

Submission archive

Download ZIP (105.71 MB)

Transparency Advocates Pushing for and Publicizing Access

Transparency advocates pushed regulators to have more information as fast as possible

Transparency too little, too late? Why and how Health Canada should make clinical data and regulatory decision-making open to scrutiny in the face of COVID-19

Sterling Edmonds, Andrea MacGregor, Agnieszka Doll, Ipek Eren Vural, Janice Graham, Katherine Fierlbeck, Joel Lexchin, Peter Doshi, Matthew Herder

Journal of Law and the Biosciences, Volume 7, Issue 1, January-June 2020, lsaa083, https://doi.org/10.1093/jlb/lsaa083

Published: 19 November 2020 **Article history**

PDF Split View Cite Permissions Share

STAT

SUBSCRIBE | LOG IN

FIRST OPINION

Far more transparency is needed for Covid-19 vaccine trials

By Jennifer E. Miller, Joseph S. Ross, and Michelle M. Mello Nov. 5, 2020

Matthew Herder @cmrher... · 6h

.@GovCanHealth Can't stop, won't stop

PUBLISHING CLINICAL data behind Covid19 interventions

$PFE & $MRNA vaccine data both up + yours to peruse within two clicks

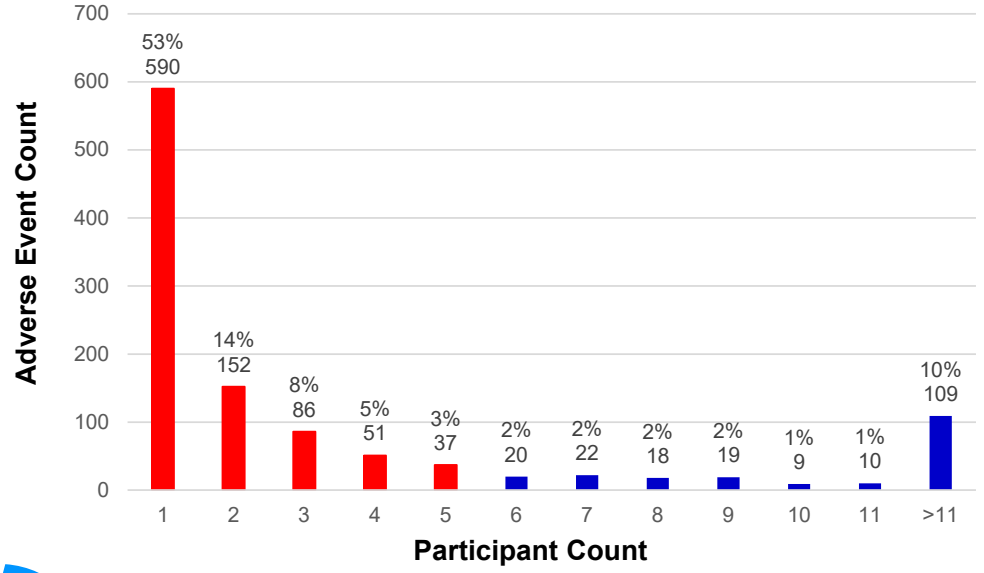
GO CANADA!

Taken from Matthew Herder Twitter post

C4591001 Preferred Term Analysis

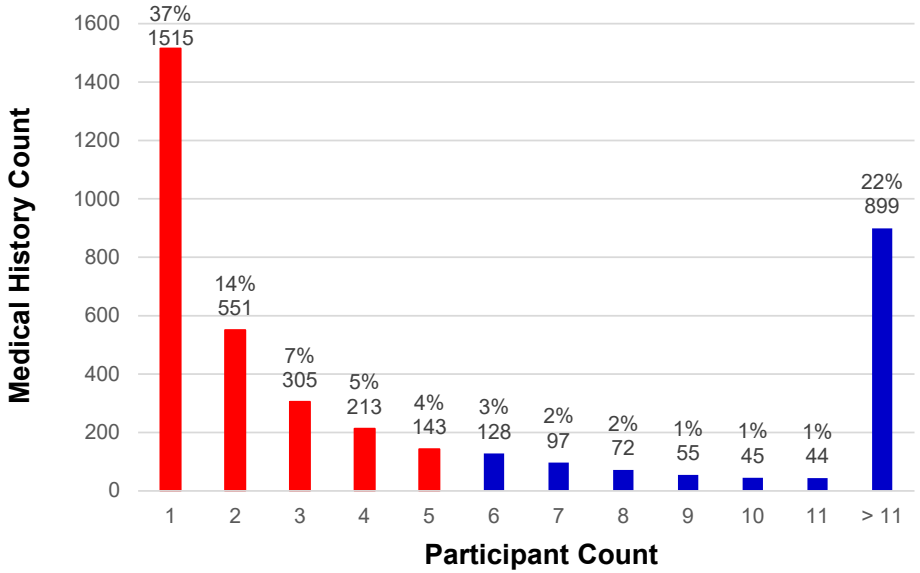
Adverse Events

- 590 (53%) of all adverse events were experienced one time by only one participant (1123 total adverse event diagnoses)
- 916 (82%) of all adverse events were experienced within the trial 5 or less times



Medical History

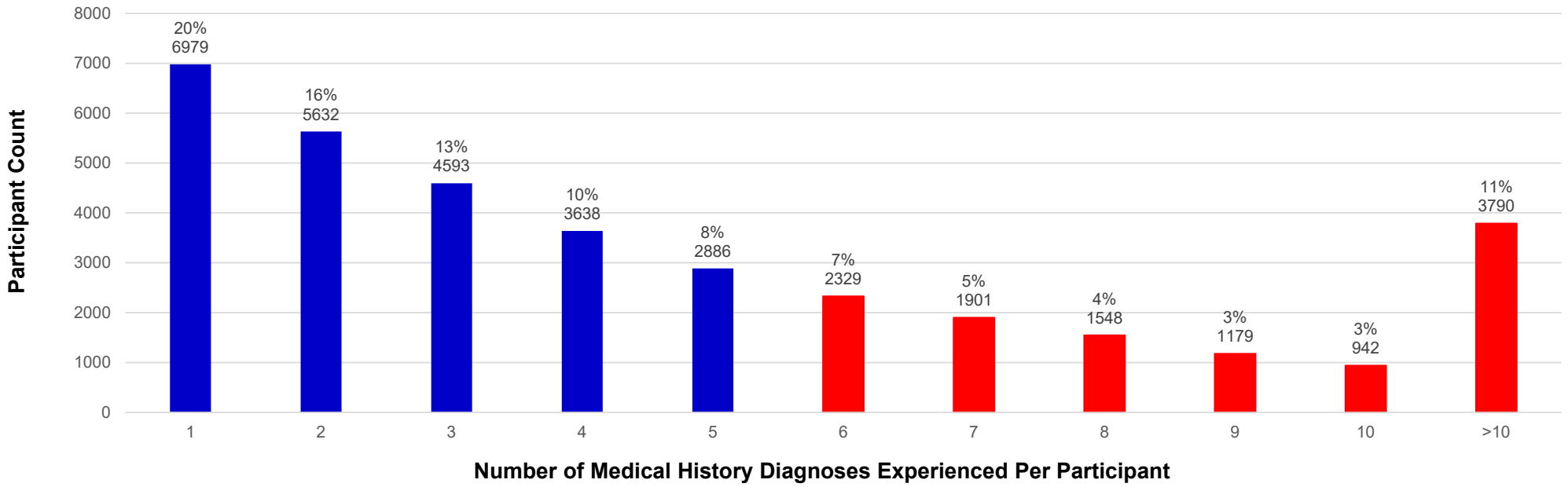
- 1515 (37%) of all medical history diagnoses were experienced one time by only one participant (4067 total medical history diagnoses)
- 2727 (67%) of all medical history diagnoses were experienced 5 or less times within the trial



C4591001 Preferred Term Analysis

Medical History Diagnoses – Count Per Participant

- 11,689 participants (33%) reported 6 or more medical history diagnoses (bars shown in red below)



C4591001 Narrative Redaction Rules

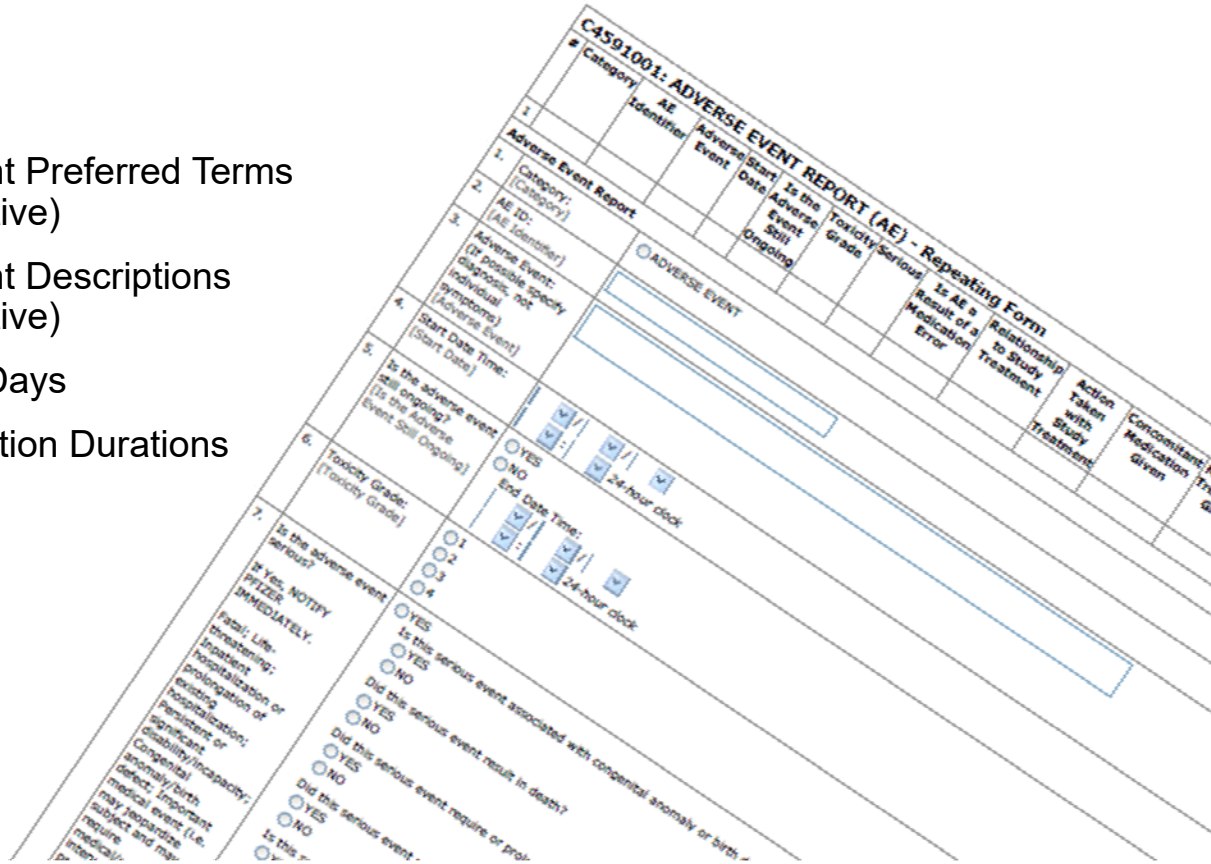
BioNTech/Pfizer agreed to strategically redact participant information, but maintained a conservative approach

Redact

- Participant ID
- Demographic Information - (Age generalized)
- Participant Dates
- Participant Locations
- Medical History
- Medical History Descriptions
- Sensitive Adverse Events:
  - Mental health diagnoses
  - Diagnoses related to reproduction
  - Sexual behavior
- Verbatim Terms

Retain

- Lab Values
- Adverse Event Preferred Terms (unless sensitive)
- Adverse Event Descriptions (unless sensitive)
- Cycle/Study Days
- Event/Medication Durations



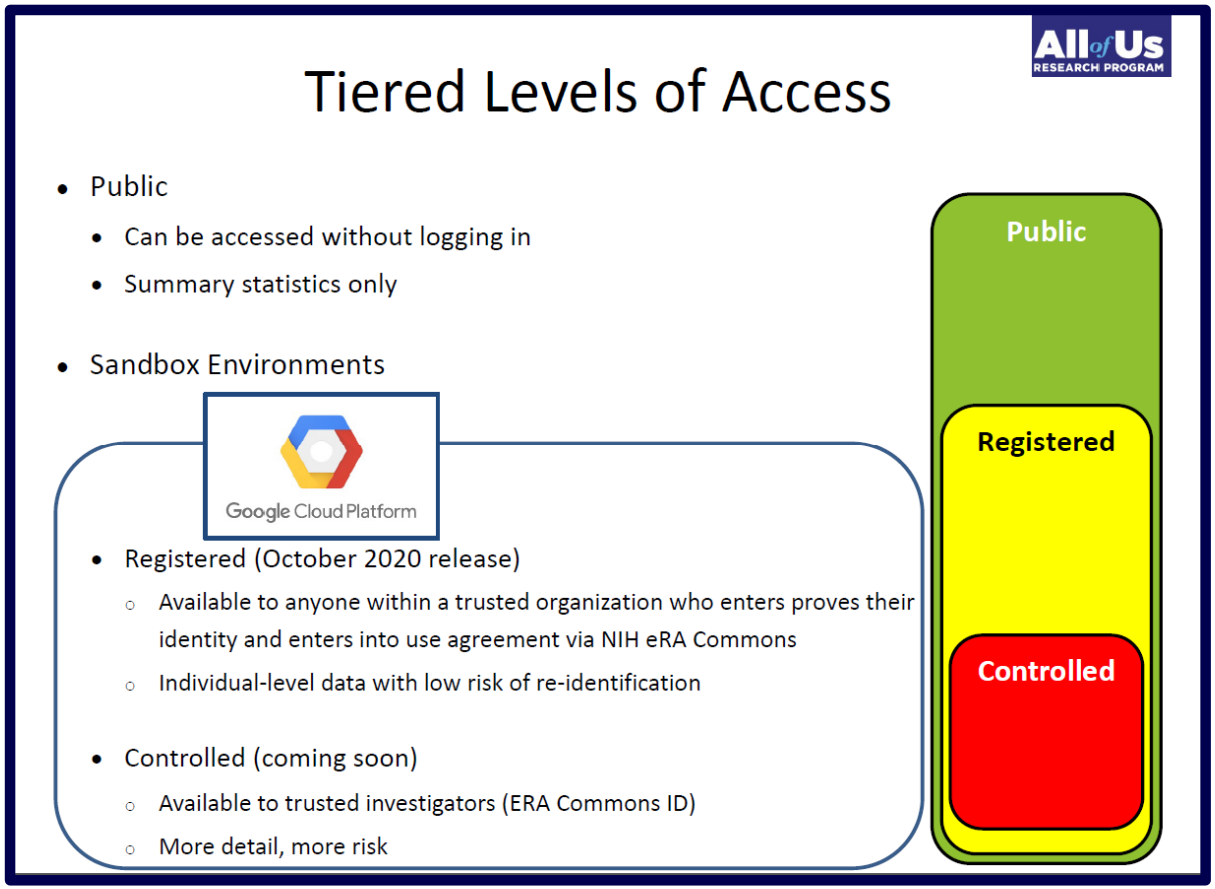
Phuse Presentation: Adversarial Modeling and Identifiability in Clinical Data Publication

Bradley Malin, Ph.D.
Vanderbilt University
Privacy Expert

June 24, 2021

NIH Database: All of Us

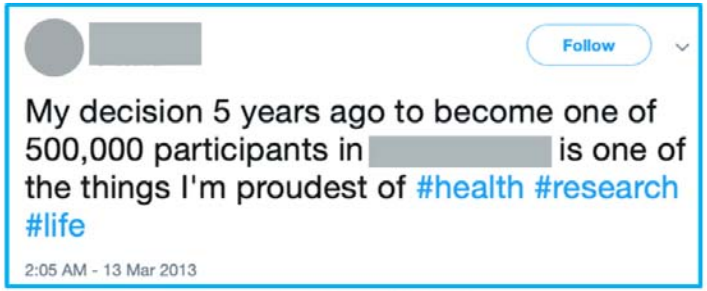
- Malin is helping to oversee the privacy and security for the "All of Us" NIH database, which includes data from:
  - Medical records
  - Survey information
  - Bio-specimens
- Options available to help protect the data beyond just anonymization:
  - Use agreements
  - Pay for Access
  - Audit
  - Unique Login/Pass
- "All of Us" database Public view:
  - Can only access summary data
  - Each summary statistic at least 20 individuals represented
- "All of Us" planned Controlled Access view:
  - Working to add additional data source types (more detailed demographics, genomic details, dates of events)
  - This group will be for individual researchers



Significant Points

Self-Disclosure is Not Helping

- Web crawling and automated classification of tweets to find participation disclosure
- Found over 100 people from 20 large cohort studies disclosing participation



Liu et al. AMIA 2019

14

What to Worry About

- The number and diversity of "investigators" will grow
- Data will be more detailed and complex
- Risk analysis helps … but we need agreement on societally acceptable levels

- Self-identification is growing issue:
  - Impact has not been fully assessed
  - They are not sharing ANY COVID data in the All of Us platform until they can better assess the risk posed by self-identification

- When assessing re-identification risk for controlled data sharing should consider behavior of rational adversaries. In contrast, with public data sharing the data steward must consider all types of adversaries to include irrational adversaries

New Malin Publication: [Protecting research data of publicly revealing participants]

Pfizer

Thank You