



d-wise™

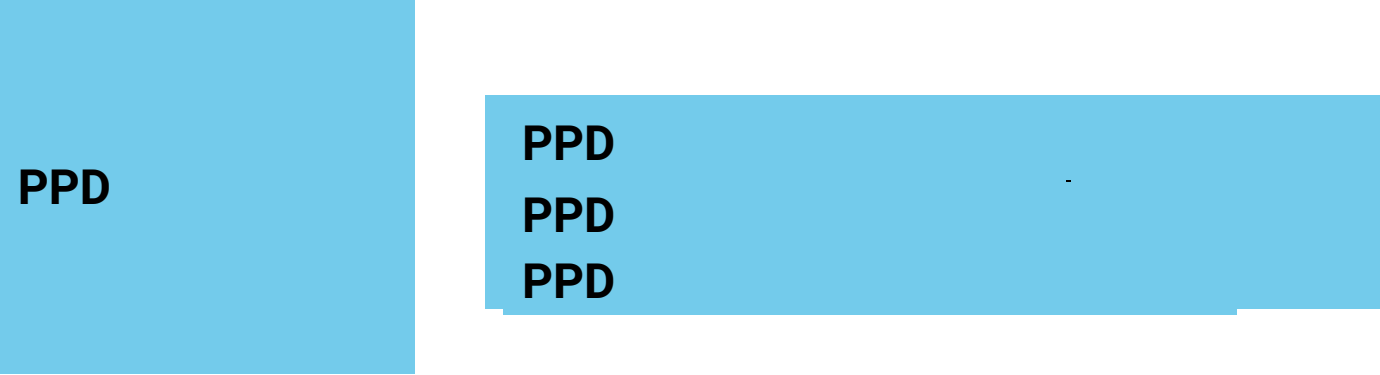
Accelerate to the Speed of Life

DH03: Achieving Balance Approaching Anonymization with Quantified Risk

Cathal Gallagher
Transparency Consultant
Cathal.Gallagher@d-wise.com

Achieving Balance

- Approaching Anonymization with Quantified risk



PPD

PPD

PPD

PPD

Agenda

- **Redaction vs Anonymisation**
- Qualitative Risk Assessment
- Quantitative risk Assessment
- Identifiers used to measure risk
- Pitfalls e.g. single deaths, partial subjid's

What is Full Redaction?

- Patient S123-101-1015 is a Black, 47 year old female from Ireland. Her first treatment date was 18th November 1996. She experienced severe headaches 2 days later on 20th November 1996. Patient died on 4th December 1996 due to suicide.

- PPD



What is Targeted Redaction?

- Patient S123-101-1015 is a Black, 47 year old female from Ireland. Her first treatment date was 18th November 1996. She experienced severe headaches 2 days later on 20th November 1996. Patient died on 4th December 1996 due to suicide.
- Patient PPD is a PPD, PPD year old PPD from PPD. PPD first treatment date was PPD PPD experienced severe headaches 2 days later on PPD. Patient died on PPD PPD due to PPD.

What is HIPS Anonymisation?

- Patient S123-101-1015 is a Black, 47 year old female from Ireland. Her first treatment date was 18th November 1996. She experienced severe headaches 2 days later on 20th November 1996. Patient died on 4th December 1996 due to suicide.
- Patient J711-818-5112 is a PPD, 41 year old female from Europe. Her first treatment date was 17th December 1996. She experienced severe headaches 2 days later on 19th December 1996. Patient died on 2nd January 1997 due to PPD.

Agenda

- Redaction vs Anonymisation
- **Qualitative Risk Assessment**
- Quantitative risk Assessment
- Identifiers used to measure risk
- Pitfalls e.g. single deaths, partial subjid's

Qualitative risk assessment

- How identifying is the information
 - Highly
 - Somewhat
 - unlikely
- How sensitive is the information
 - Unsensitive
 - Sensitive

Qualitative risk assessment

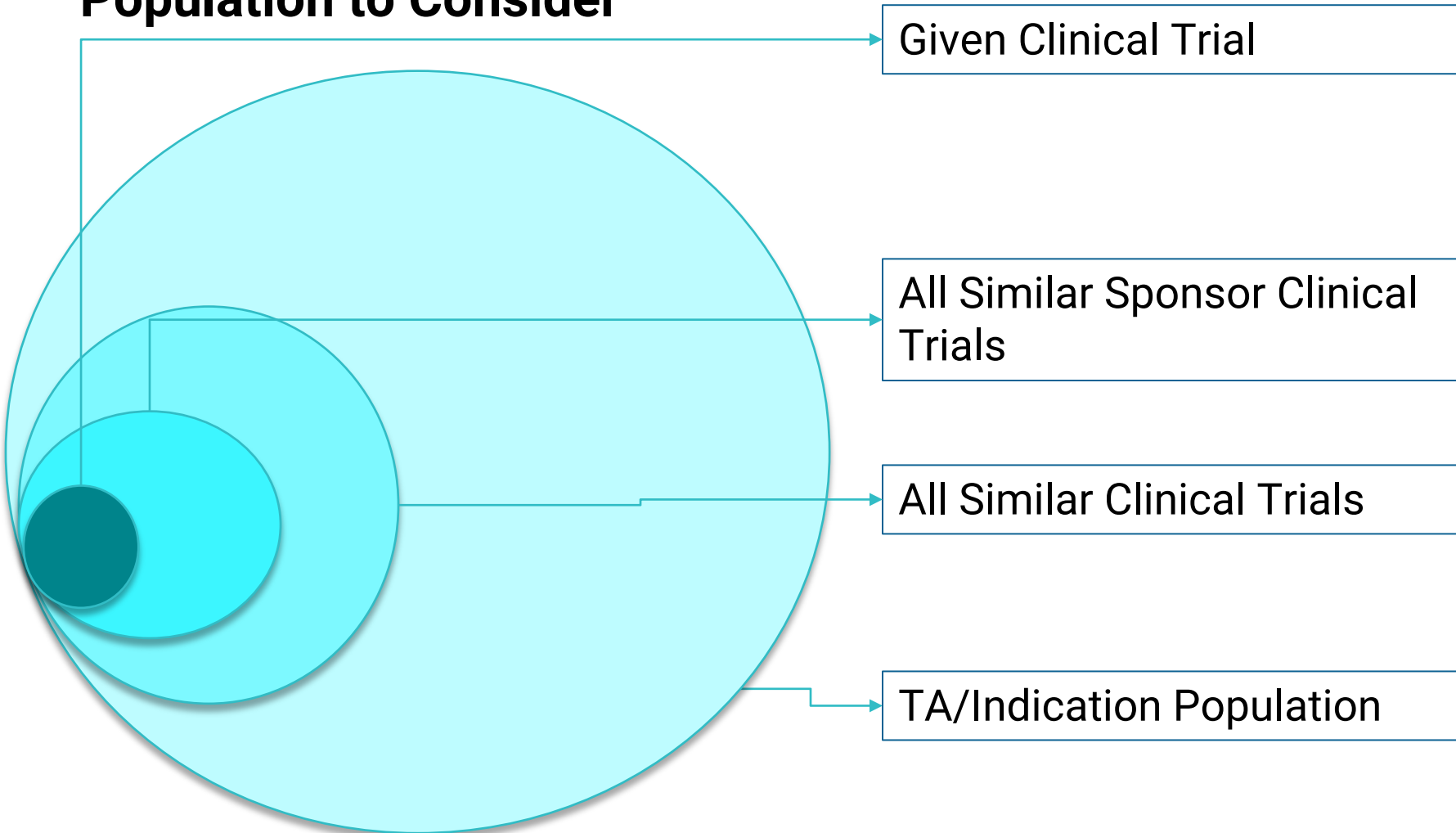
How Identifying	How Sensitive	Overall Risk
Unlikely	Unsensitive	Low
Somewhat	Unsensitive	Medium
Highly	Unsensitive	High
Unlikely	Sensitive	Medium
Somewhat	Sensitive	High
Highly	Sensitive	High

Agenda

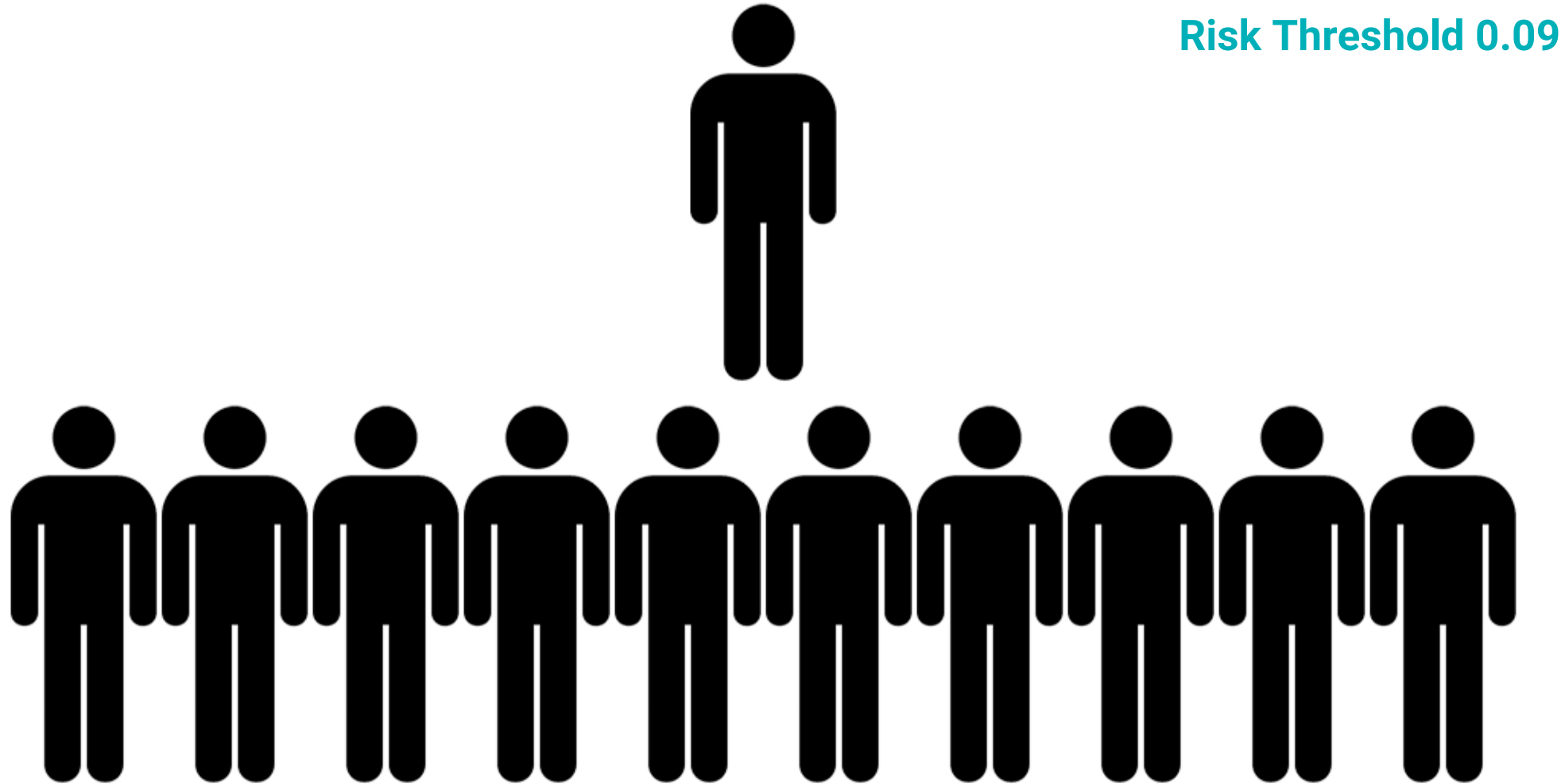
- Redaction vs Anonymisation
- Qualitative Risk Assessment
- **Quantitative risk Assessment**
- Identifiers used to measure risk
- Pitfalls e.g. single deaths, partial subjid's

What is your Population Size?

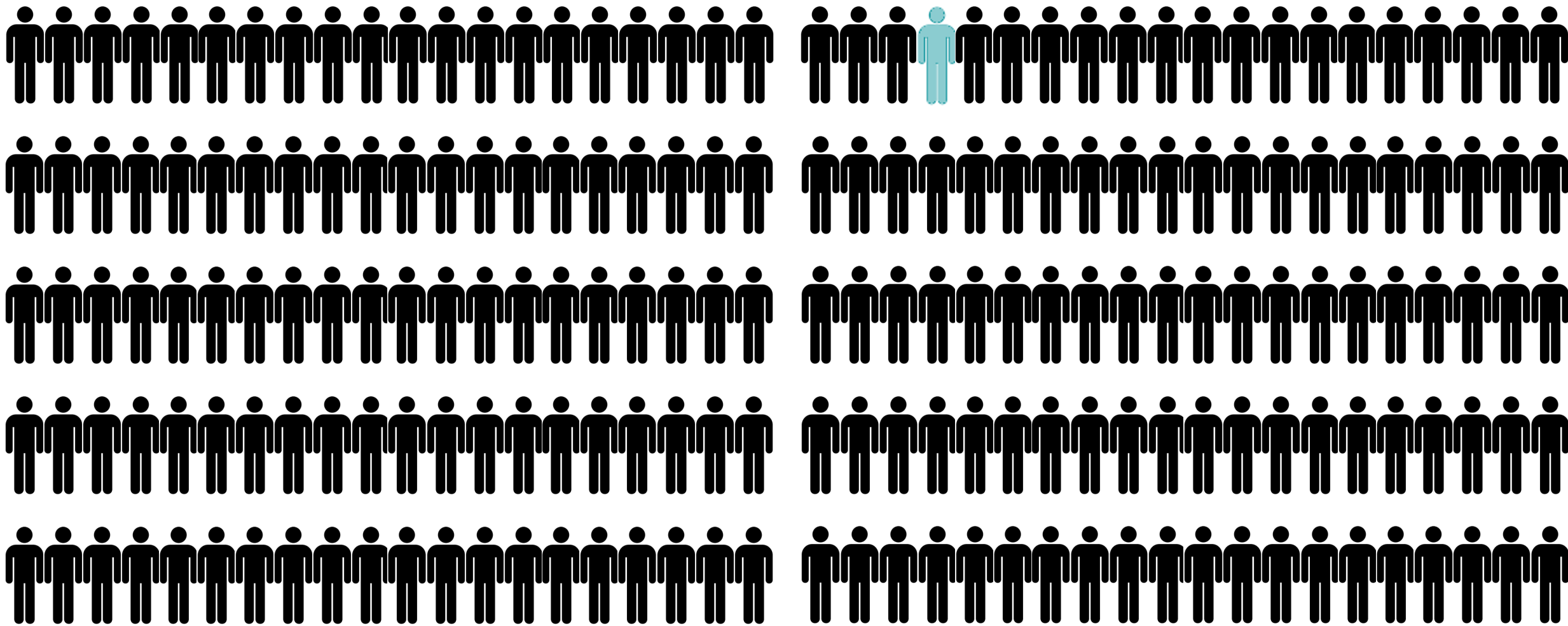
Population to Consider



Quantitative risk assessment



$$1/200 = .005$$



Agenda

- Redaction vs Anonymisation
- Qualitative Risk Assessment
- **Quantitative risk Assessment**
- **Identifiers used to measure risk**
- Pitfalls e.g. single deaths, partial subjid's

What is an identifier?

What is an identifier?



Sex



Height



Weight



Age



Country

- White
- Black
- Asian
- Pacific Islander

Race

- Hispanic
- Non-Hispanic

Ethnicity

- 17.4
- 19.9
- 27.2
- 35.4

BMI



Sex

- Keep
- Redact



Height

- Keep
- 5 Integer Band
- 10 Integer Band
- 20 Integer Band
- 25 Integer Band
- 30 Integer Band
- 40 Integer Band
- 50 Integer Band
- Redact



Weight

- Keep
- 5 Integer Band
- 10 Integer Band
- 20 Integer Band
- 25 Integer Band
- 30 Integer Band
- 40 Integer Band
- 50 Integer Band
- Redact



Age

- Keep
- 5 Integer Band
- 10 Integer Band
- 20 Integer Band
- 25 Integer Band
- 30 Integer Band
- Child/Adult
- Redact



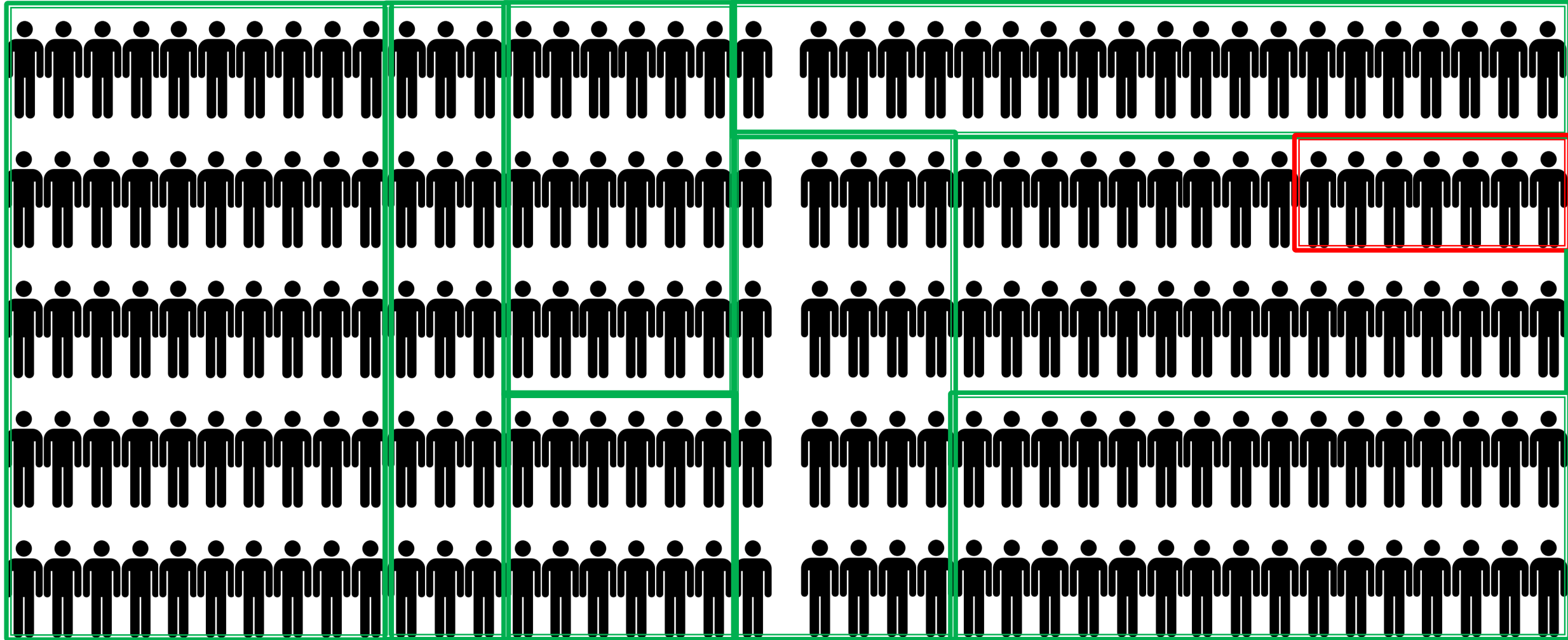
Country



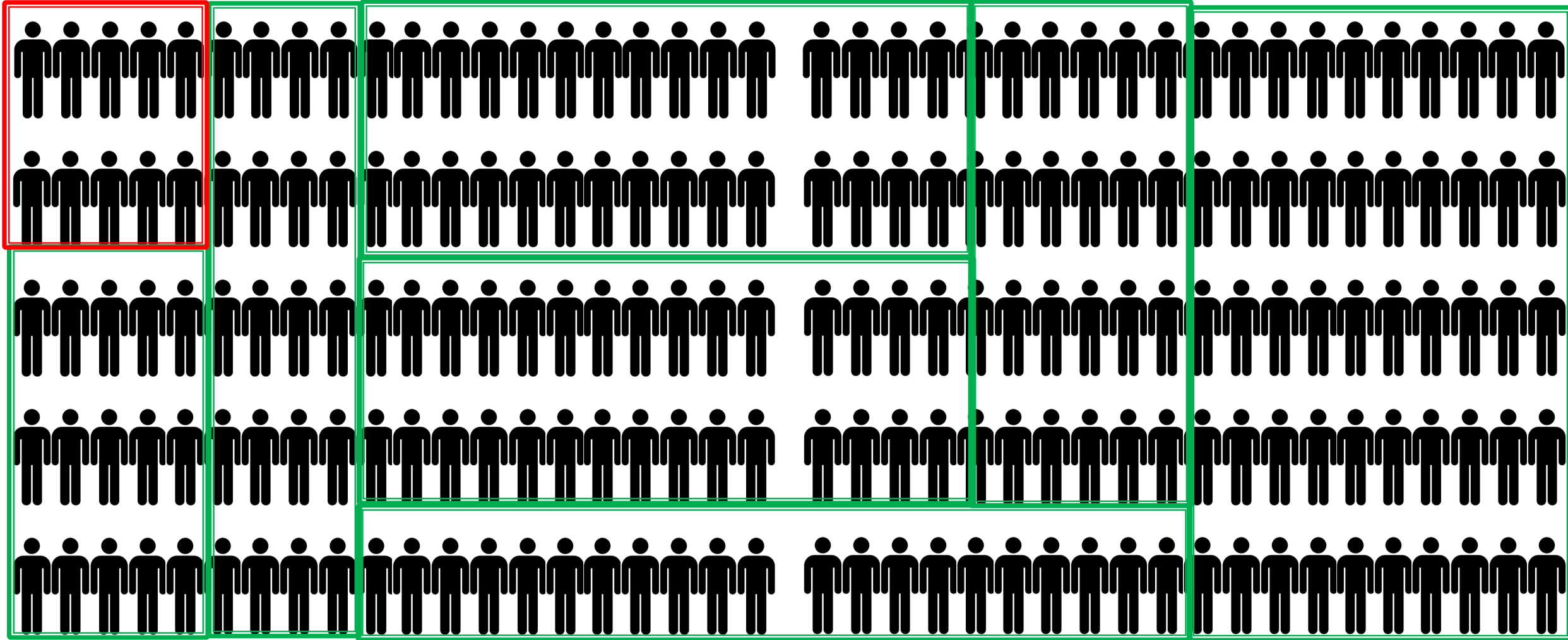
- Keep
- Low Frequency(11)
- Low Frequency(20)
- Low Frequency(30)
- Low Frequency(40)
- Redact

$2 \times 9 \times 9 \times 8 \times 6 =$
7776 combinations

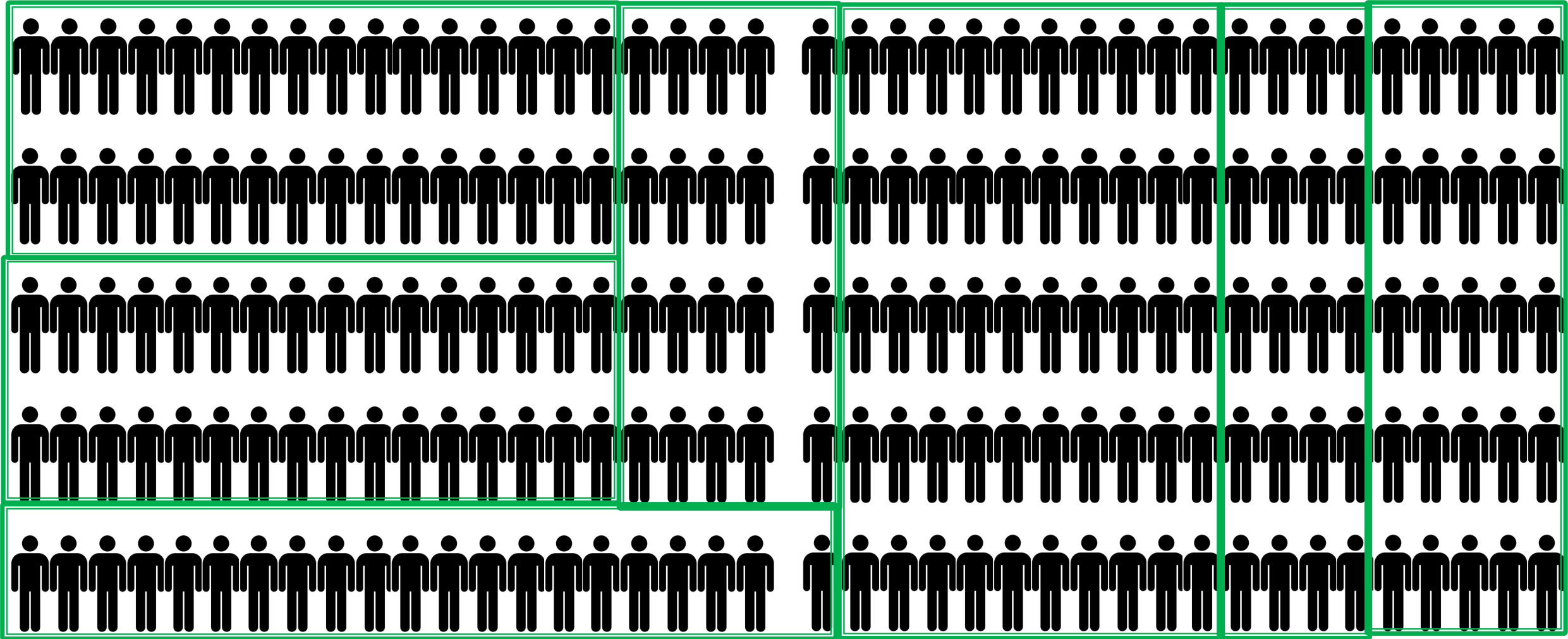
$$1/7 = .143$$



$$1/10 = .1$$



$$1/20 = .05$$



SubjID	Sex	County	Age
1001	M	USA	21
1002	F	Spain	22
1003	M	USA	28
1004	F	USA	34
1005	F	Ireland	39
1006	F	USA	21
1007	M	USA	33
1008	M	Spain	34
1009	M	Ireland	25
1010	F	Ireland	25
1011	M	Ireland	26
1012	F	Spain	38
1013	M	Spain	37
1014	F	USA	37
1015	F	USA	22
1016	M	Spain	20
1017	M	Ireland	30
1018	F	Ireland	29
1019	M	USA	29
1020	F	Ireland	21
1021	M	Spain	27
1022	M	Spain	38
1023	M	Ireland	33
1024	M	USA	36
1025	F	USA	26

SubjID	Sex	County	Age
1005	F	Ireland	30 to 39
1017	M	Ireland	30 to 39
1023	M	Ireland	30 to 39
1012	F	Spain	30 to 39
1008	M	Spain	30 to 39
1013	M	Spain	30 to 39
1022	M	Spain	30 to 39
1004	F	USA	30 to 39
1014	F	USA	30 to 39
1007	M	USA	30 to 39
1024	M	USA	30 to 39
1010	F	Ireland	20 to 29
1018	F	Ireland	20 to 29
1020	F	Ireland	20 to 29
1009	M	Ireland	20 to 29
1011	M	Ireland	20 to 29
1002	F	Spain	20 to 29
1016	M	Spain	20 to 29
1021	M	Spain	20 to 29
1006	F	USA	20 to 29
1015	F	USA	20 to 29
1025	F	USA	20 to 29
1001	M	USA	20 to 29
1003	M	USA	20 to 29
1019	M	USA	20 to 29

SubjID	Sex	County	Age
1010	F	Europe	20 to 29
1018	F	Europe	20 to 29
1020	F	Europe	20 to 29
1002	F	Europe	20 to 29
1009	M	Europe	20 to 29
1011	M	Europe	20 to 29
1016	M	Europe	20 to 29
1021	M	Europe	20 to 29
1006	F	USA	20 to 29
1015	F	USA	20 to 29
1025	F	USA	20 to 29
1001	M	USA	20 to 29
1003	M	USA	20 to 29
1019	M	USA	20 to 29
1005	F	Europe	30 to 39
1012	F	Europe	30 to 39
1017	M	Europe	30 to 39
1023	M	Europe	30 to 39
1008	M	Europe	30 to 39
1013	M	Europe	30 to 39
1022	M	Europe	30 to 39
1004	F	USA	30 to 39
1014	F	USA	30 to 39
1007	M	USA	30 to 39
1024	M	USA	30 to 39

Anonymisation

Redact

Child/Adult Band

30 Integer Band

~~25 Integer Band~~

20 Integer Band

10 Integer Band

5 Integer Band

Keep



Data Utility

Redact

Child/Adult Band

30 Integer Band

~~25 Integer Band~~

20 Integer Band

10 Integer Band

5 Integer Band

Keep

Agenda

- Redaction vs Anonymisation
- Qualitative Risk Assessment
- Quantitative risk assessment
- Identifiers used to measure risk
- **Pitfalls e.g. single deaths, partial subjid's**

Pit falls: Single Deaths

- Patient S123-101-1015 is a Black, 47 year old female from Ireland. Her first treatment date was 18th November 1996. She experienced severe headaches 2 days later on 20th November 1996. Patient died on 4th December 1996 due to suicide.
- Patient J711-818-5112 is a PPD, 41 year old female from Europe. Her first treatment date was 17th December 1996. She experienced severe headaches 2 days later on 19th December 1996. Patient died on 2nd January 1997 due to PPD.

Pit falls: Single Deaths

- A total of 98 (92.5%) patients were continuing in the study and 8 (7.5%) subjects had discontinued from the study. The primary reason for study discontinuation was pregnancy (6 patients, 5.6%) 1 patient died due to suicide.
- Patient J711-818-5112 is a PPD, 41 year old female from Europe. Her first treatment date was 17th December 1996. She experienced severe headaches 2 days later on 19th December 1996. Patient died on 2nd January 1997 due to PPD.

Pit falls: Partial Subject IDs

- Patient S123-101-1015 is a Black, 47 year old female from Ireland. Her first treatment date was 18th November 1996. She experienced severe headaches 2 days later on 20th November 1996. Patient died on 4th December 1996 due to suicide.
- Patient J711-818-5112 is a PPD, 41 year old female from Europe. Her first treatment date was 17th December 1996. She experienced severe headaches 2 days later on 19th December 1996. Patient died on 2nd January 1997 due to PPD.

Pit falls: Partial Subject IDs

- A total of 98 (92.5%) patients were continuing in the study and 8 (7.5%) subjects had discontinued from the study. The primary reason for study discontinuation was pregnancy (6 patients, 5.6%) 1 patient died due to suicide(1015).
- A total of 98 (92.5%) patients were continuing in the study and 8 (7.5%) subjects had discontinued from the study. The primary reason for study discontinuation was pregnancy (6 patients, 5.6%) 1 patient died due to PPD (5112).

Questions?

Cathal.Gallagher@d-wise.com

