

# Analysis of CSRs already published (PhUSE review)

Lukasz Kniola, Biogen, UK

EMA Technical Anonymisation Group (TAG) Meeting

29-30 November 2017

# Published CSRs Review Spreadsheet



Prepared by  
PhUSE De-Identification  
Working Group

Will be publically  
available, searchable.

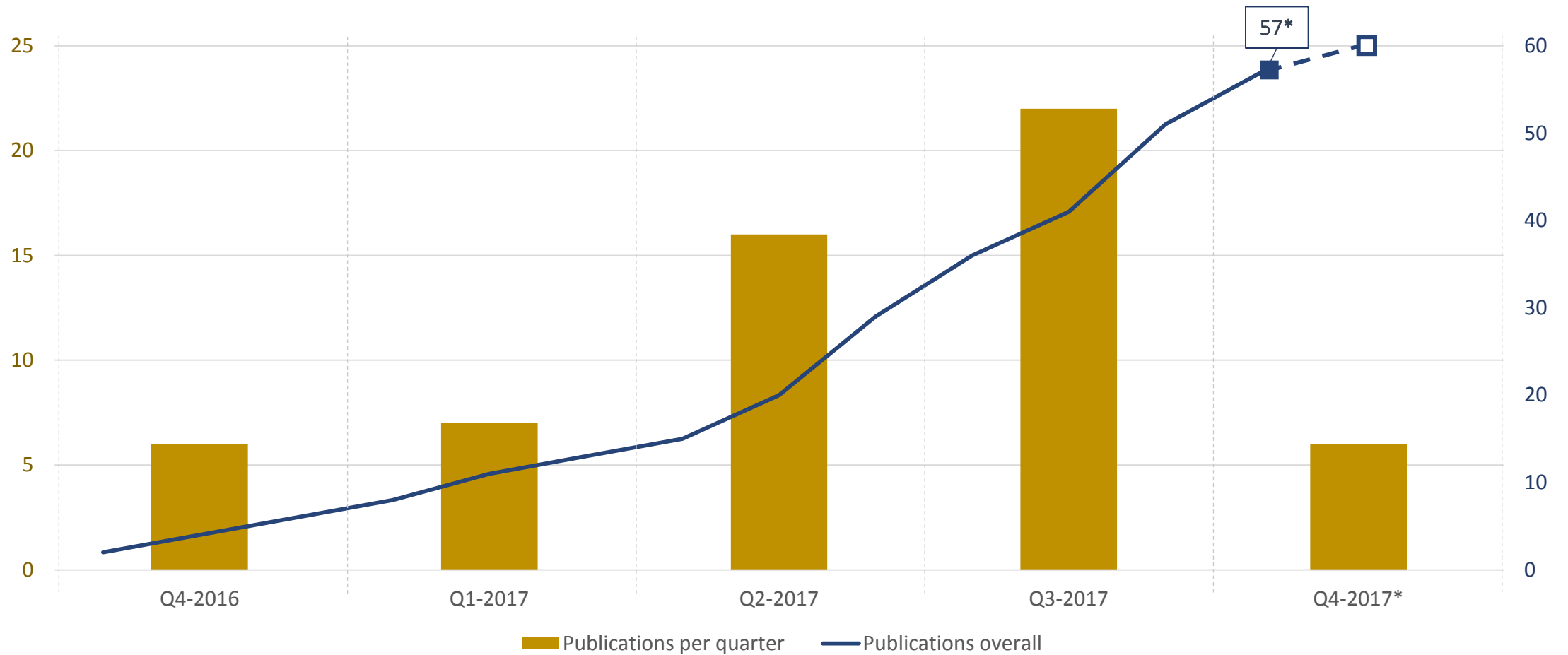
Filters to simplify  
searches and find similar  
publications

Published CSRs Review															
ITG Isotope Technologies Garching GmbH															
	A	B	C	D	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	First Review	Second Review	Comments	General	Number of Studies with <100 Subjects	Number of Studies with 100-1000 Subjects	Number of Studies with 1000+ Subjects	Rules							
				Product Name				Anonymisation method	Narratives	Listings	Subject ID	Dates	Assessment methodology		Quasi
2	LK/ BH	BX		Kyprolis				Redaction	Fully redacted	Redacted	Redaction	Redaction	Qualitative		
3	LK/ BH	CC	narrative redaction	Zurampic				Redaction	Partially redacted	Combination	Redaction	Redaction	Qualitative		Sex, a
4	LK/ BH	CC	clarify generic status	Armisarte				N/A	N/A	N/A	N/A	N/A	N/A		
5	LK/ BH	CC		Caspofungin Accord				N/A	N/A	N/A	N/A	N/A	N/A		
6	LK/ BH	CC	CC: added list of retained quasi identifiers	Praxbind				Redaction	Fully redacted	Combination	Redaction	Redaction	Qualitative		Type: e.g. large inform larger studies comb quasi
7	LK/ BH	CC	CC: added list of de-identified quasi identifiers	Tarceva				Redaction	Fully redacted	Redacted	Redaction	Redaction	Qualitative		
8	LK/ BH	CC		Palonosetron Hospira				N/A	N/A	N/A	N/A	N/A	N/A		
9	LK/ BH			Aripiprazole Mylan				Redaction			Redaction	Redaction	Fulfilment of 3 anonymisa		
10	LK/ BH			Cubicin				Redaction	Fully redacted	Redacted	Redaction	Redaction	Qualitative		
11				Coagadex				Redaction	Fully redacted	Redacted	Redaction	Redaction	Fulfilment of 3 anonymisa		
12	LK/ BH			Empliciti				Redaction	Partially redacted	Redacted	Redaction	Redaction	Qualitative		

# Categories

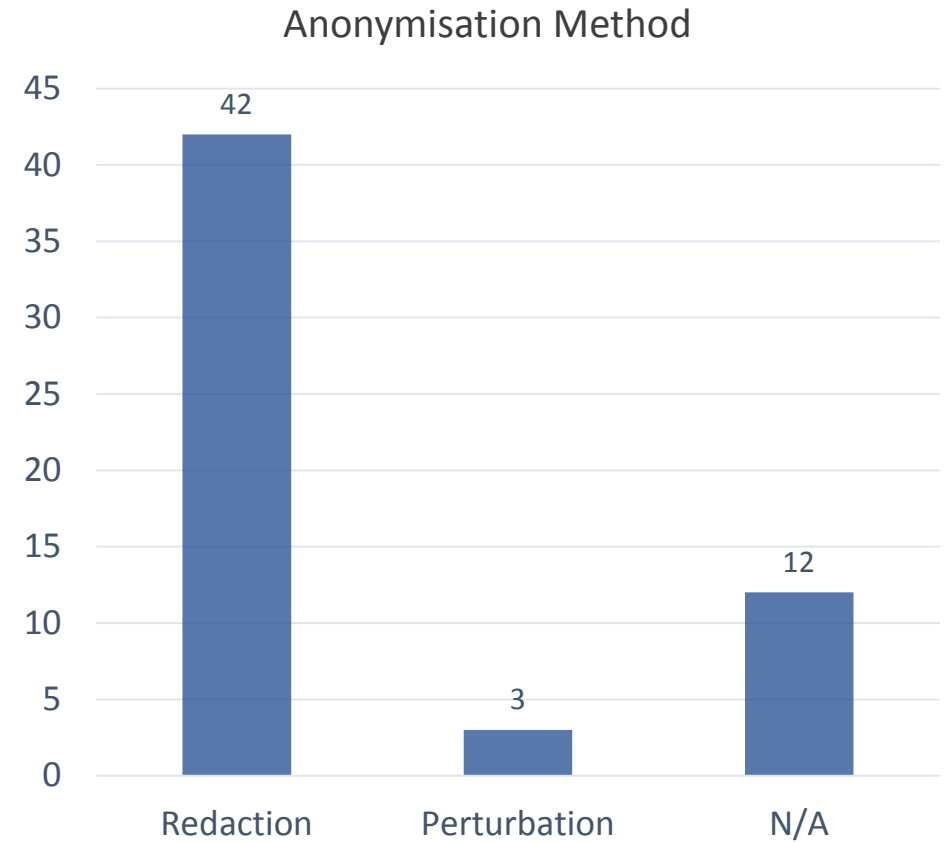
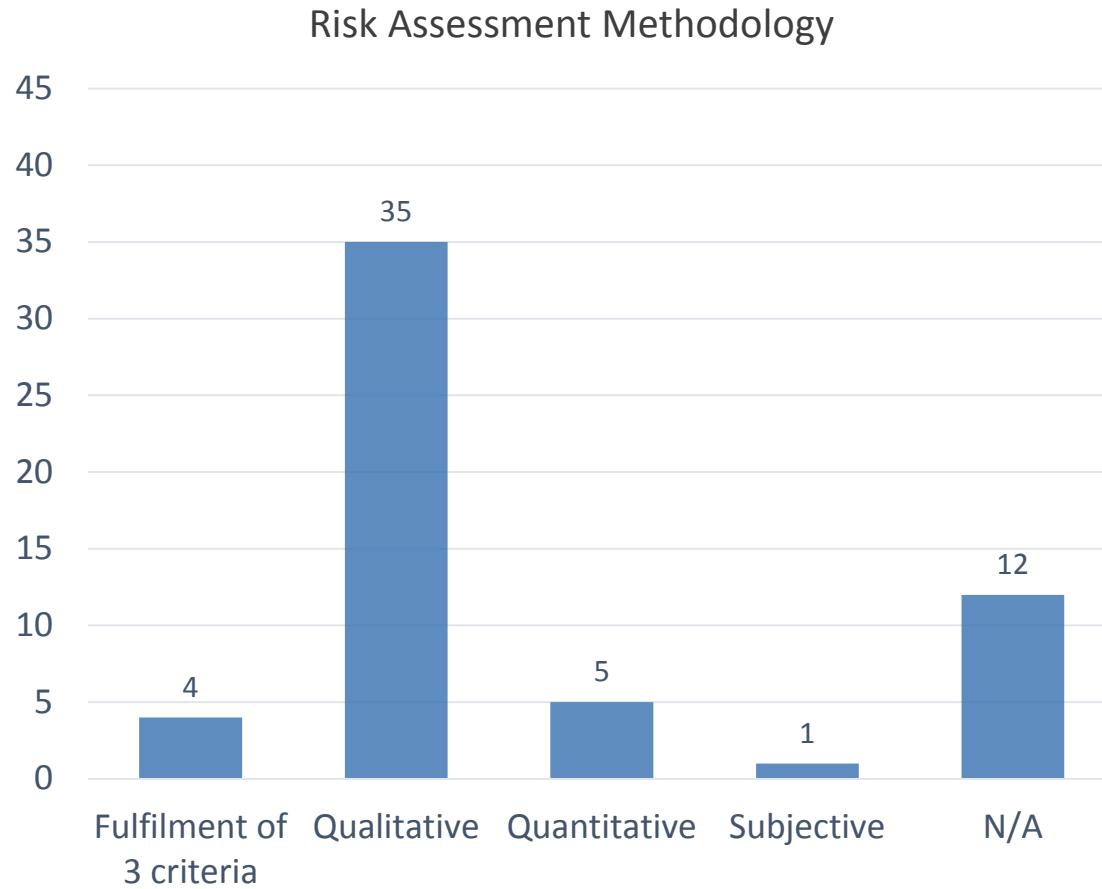
EMA admin	Size	Rules	Risk Assessment	Other
Product Name	Number of Documents	Anonymisation method	Assessment methodology	Data utility considerations
Active Substance	Number of Studies Overall	Narratives	Quasi-identifiers: retained	Special provision for sensitive information
Sponsor	Number of Studies with <100 Subjects	In-text Listings	Quasi-identifiers: de-identified	Special provision for non-subject personal data (vendors, PIs, etc.)
Product Status	Number of Studies with 100-<1000 Subjects	Subject ID	Population used for risk calculation	
Publication Date	Number of Studies with 1000+ Subjects	Dates	Threshold used	
Procedure Type				
AtcCode				
Generic				
Biosimilar				
Procedure Number				
etc.				

# Publications Over Time

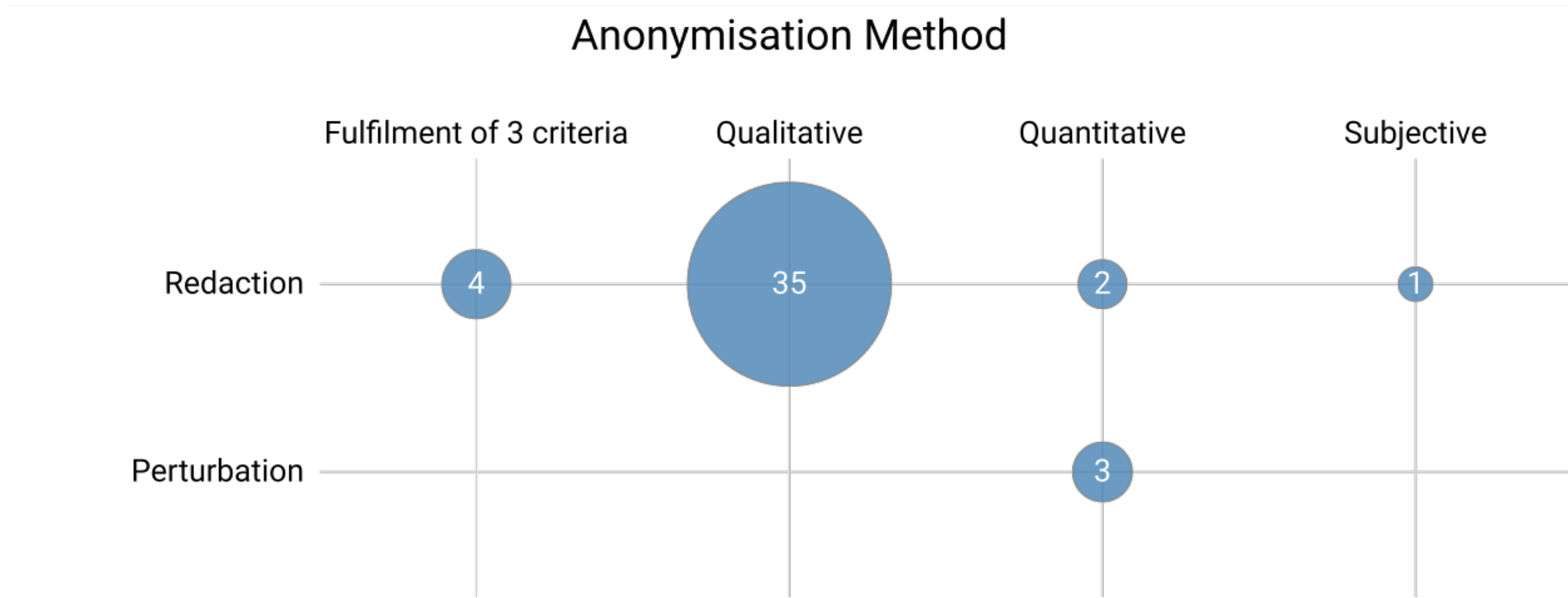


\*) As of 31-Oct-2017

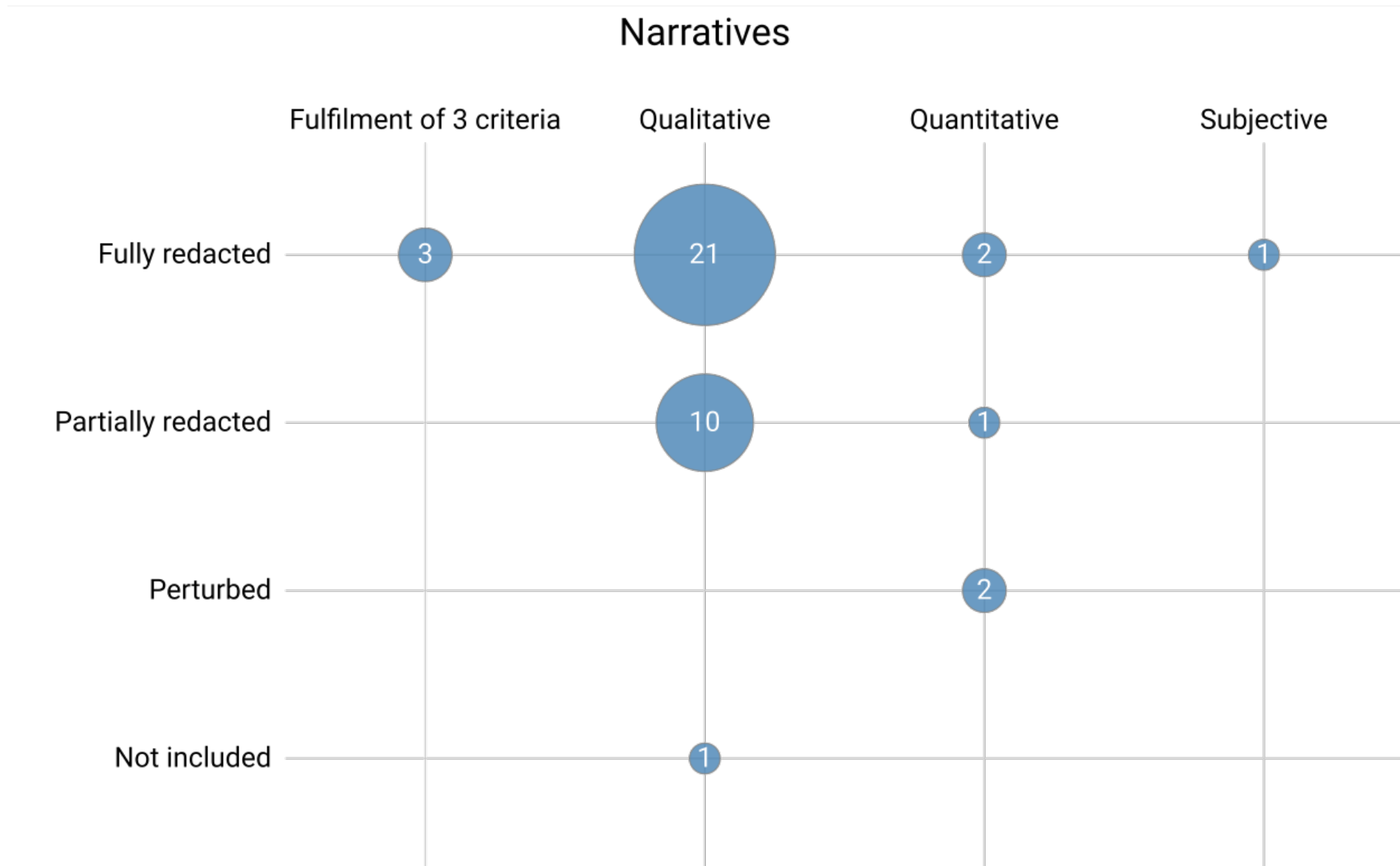
# Summary of findings



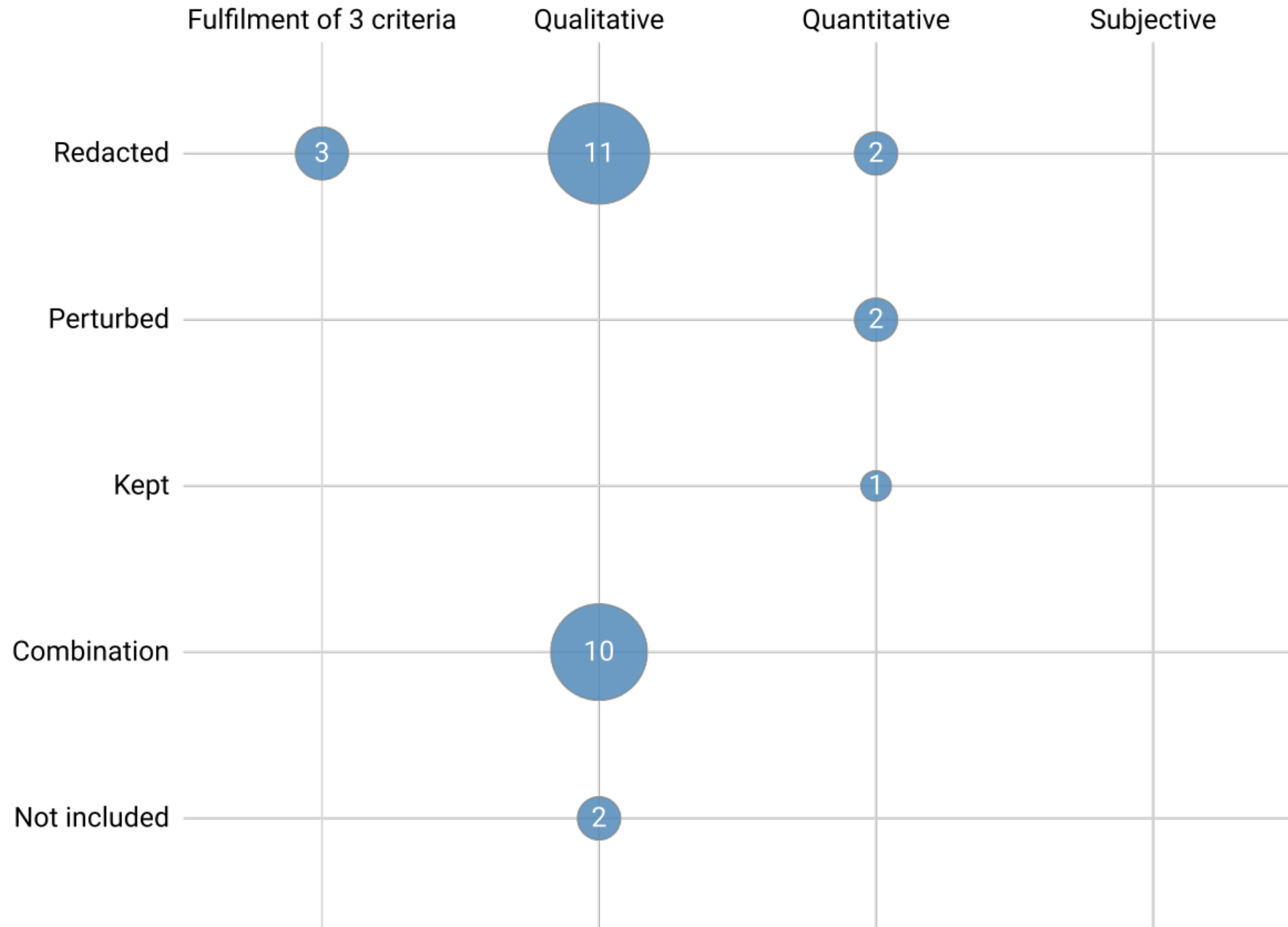
# Summary of findings – cont.



# Summary of findings – cont.

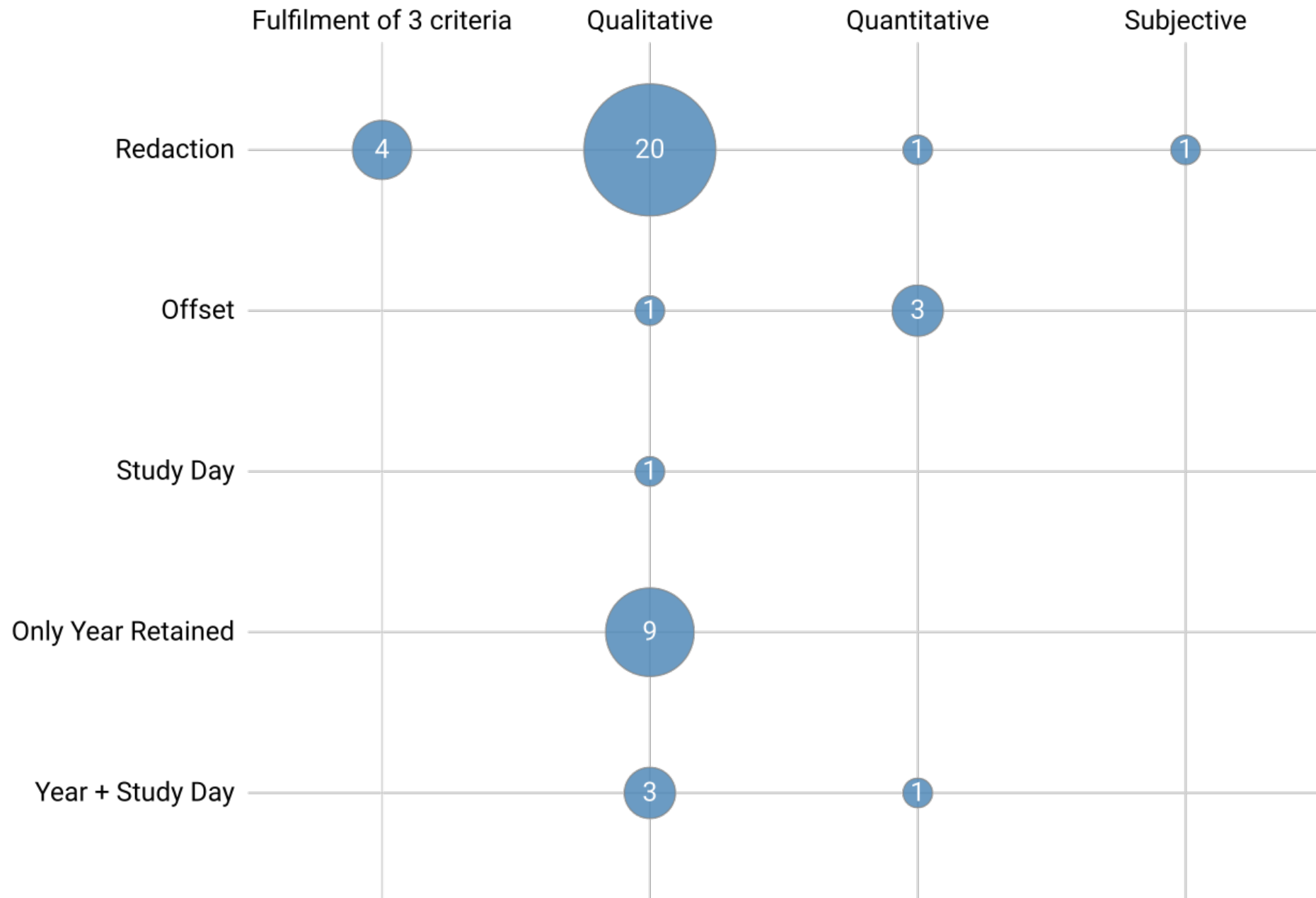


## In-text Listings





# Summary of findings – cont.



# Common Quasi-Identifiers

- Age, sex, race/ethnicity, country, site
  - Height, weight, BMA, BSA
  - Medical History
  - Dates
- 
- SAEs, ConMeds, Labs

# Quantitative Approach - details

General				Rules								
Product Name	Sponsor	Orphan	Number of Studies (<100/ 100-<1000/ 1000+ subjects)	Anonymisation method	Narratives	Listings	Subject ID	Dates	Quasi-identifiers: retained	Quasi-identifiers: de-identified	Population used for risk calculation	Threshold used
Alprolix	Swedish Orphan Biovitrum AB	Yes	5 (3/2/0)	Redaction	Fully redacted	Redacted	Redaction	Redaction	None	Age, race, country, height, weight, BMI, SAEs, surgeries, bleeding episodes	Submission population	0.09
Zinbryta	Biogen Idec Ltd	No	11 (4/6/1)	Redaction	Partially redacted	Kept	Redaction	Year & Study Day	Sex, year of dates, coded terms	Age, race, country, profession, height, weight	Submission population	0.09
Afinitor	Novartis Europharm Ltd	No	1 (0/1/0)	Perturbation	Fully redacted	Redacted	Perturbation	Offset	QIs perturbed	Gender, age, race, site ID, date, MH, SAE	Similar studies, any sponsor	0.09
Darzalex	Janssen-Cilag International NV	Yes	5 (4/1/0)	Perturbation	Perturbed	Perturbed	Perturbation	Offset	QIs perturbed	Gender, age, race, site ID, date, MH, SAE, height, weight, BMI, drinking, smoking habits	Similar studies, any sponsor	0.09
Imbruvica	Janssen-Cilag International NV	Yes	3 (0/3/0)	Perturbation	Perturbed	Perturbed	Perturbation	Offset	QIs perturbed	Gender, age, race, site ID, date, MH, SAE	Similar studies, any sponsor	0.09

# Data utility considerations

- Aggregate summaries and analyses have the most scientific value and were kept.
- Patient-level data redacted due to:
  - Rare disease
  - Sensitive information
  - Unique information
  - Unfamiliar process
  - Presumed very detailed knowledge by plausible adversaries

# Questions and Discussion