

Society for Clinical Trials 2016

Distribution of Item Types Used in Case Report Forms Implemented in Modern Data Management Systems

Andie Lears

Center for Clinical Trials
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

18-May-2016
Montréal, Quebec

Prior to 2001

- Multi-center RCTs at the CCT used distributed data systems
 - Disc-based on stand-alone PCs
 - Pure character input via SAS programs
 - E.g. typing “1” for Yes, “2” for No
 - Requires extra programming to set limits on values
 - Periodic data harvest via mailed discs

Modern Data Management Systems

- Starting in 2001, in-house development of web-based system based on paper CRFs
 - Avoid character entry when possible
 - Modern interface elements
 - E.g. radio buttons, checkboxes, dropdown lists, etc.
 - Limits are set by the options provided (e.g. “Yes” or “No” are only choices)
- Since 2001, used in approximately 3 dozen completed and ongoing RCTs and other studies

Analysis

- Analyzed all data elements collected for several recent completed studies in 2 of our networks from 2004-2015
 - Asthma (5 trials)
 - Alzheimer's (2 trials)
- Total of 7 trials
 - 171 forms (24/study)
 - 9264 items (37/form median)
 - 226,894 records (463/form median)
 - 14,242,577 total items entered

Single choice items

- Yes/No

16. Do you have allergies that make your asthma worse:

(^{Yes}
1) (^{No}
2)

16. Do you have allergies that make your asthma worse:

(1) Yes (2) No

- Single choice

7. Gender (*check only one*):

Male (¹)

Female (²)

7. Gender:
(*check only one*)

(1) Male
 (2) Female

Multiple choice items

11. Type of unusual event (*check all that apply*)

a. Protocol deviation: (1)

b. Unanticipated medical problem: (1)

c. Unusual problem: (1)

d. Other (*specify*): (1)

_____ type of event

11. Type of unusual event:
(*check all that apply*)

(a) Protocol deviation:

(b) Unanticipated medical problem:

(c) Unusual problem:

(d) Other (*specify*):

Numeric items

11. Age:

____ _
years

11. Age:

Date items

9. Date of event onset:

____ | ____ | ____ | ____ | ____ | ____
day | mon | year

9. Date of event onset:

Text items

9. Event (*specify*):

name event

9. Event (*specify*):

Distribution of item types

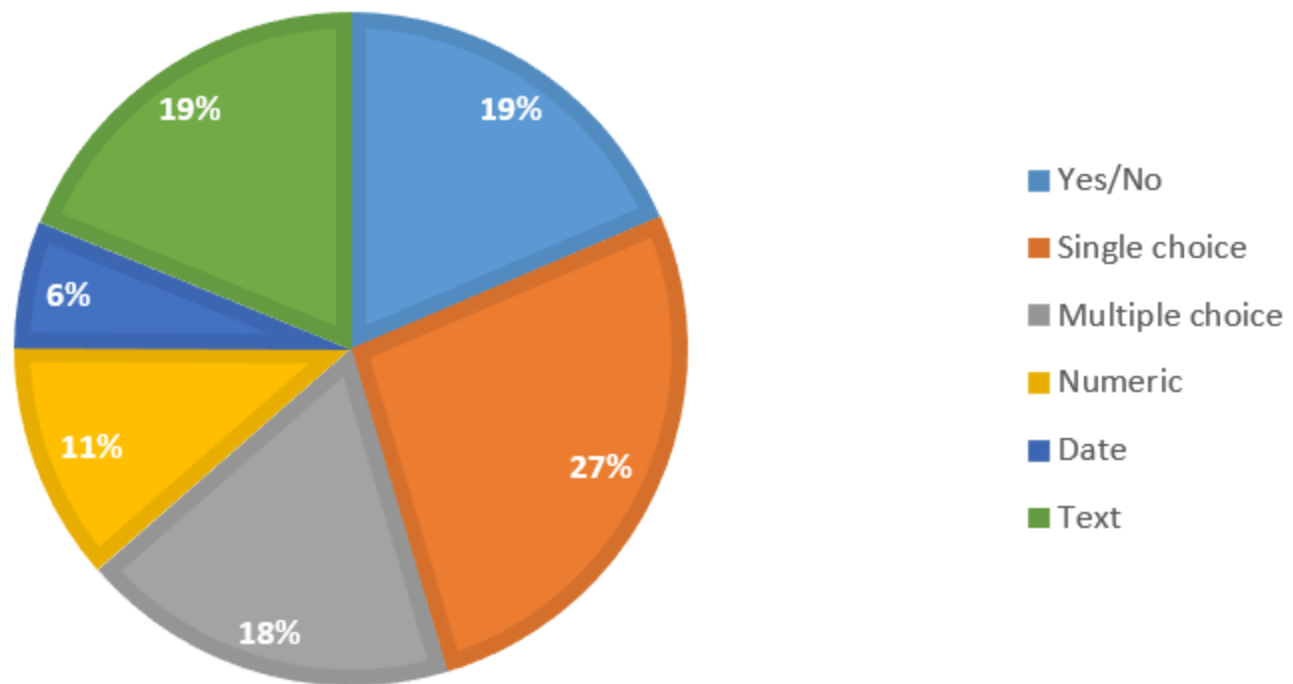
	#	%
Yes/No	1719	19
Single choice	2484	27
Multiple choice	1688	18
Numeric	1063	11
Date	566	6
Text	1744	19
Total	9264	

Distribution of item types by network

	Alzheimer's		Asthma	
	#	%	#	%
Yes/No	387	17	1332	19
Single choice	733	32	1751	25
Multiple choice	401	18	1287	18
Numeric	128	6	935	13
Date	168	7	398	6
Text	450	20	1294	18
Total	2267		6997	

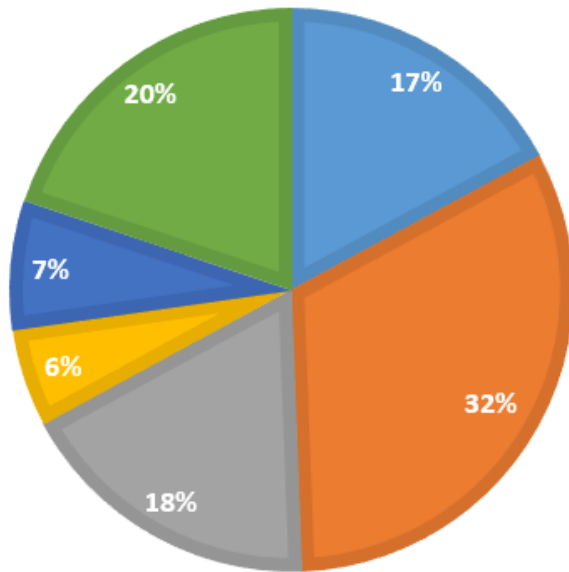
Distribution of item types

DISTRIBUTION OF ITEM TYPES

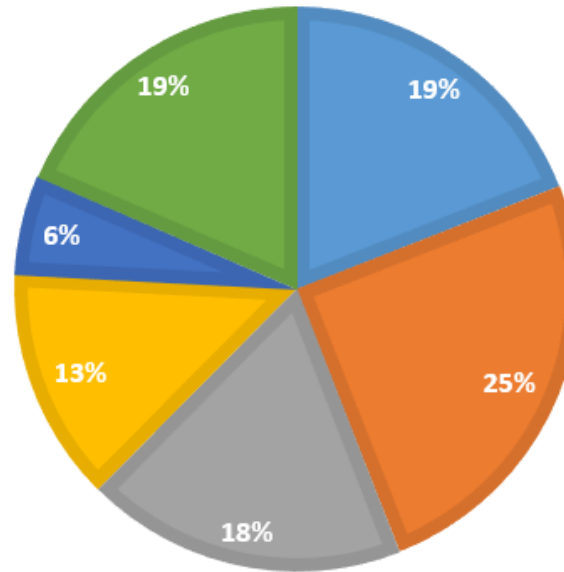


Distribution of item types by network

ALZHEIMER'S



ASTHMA



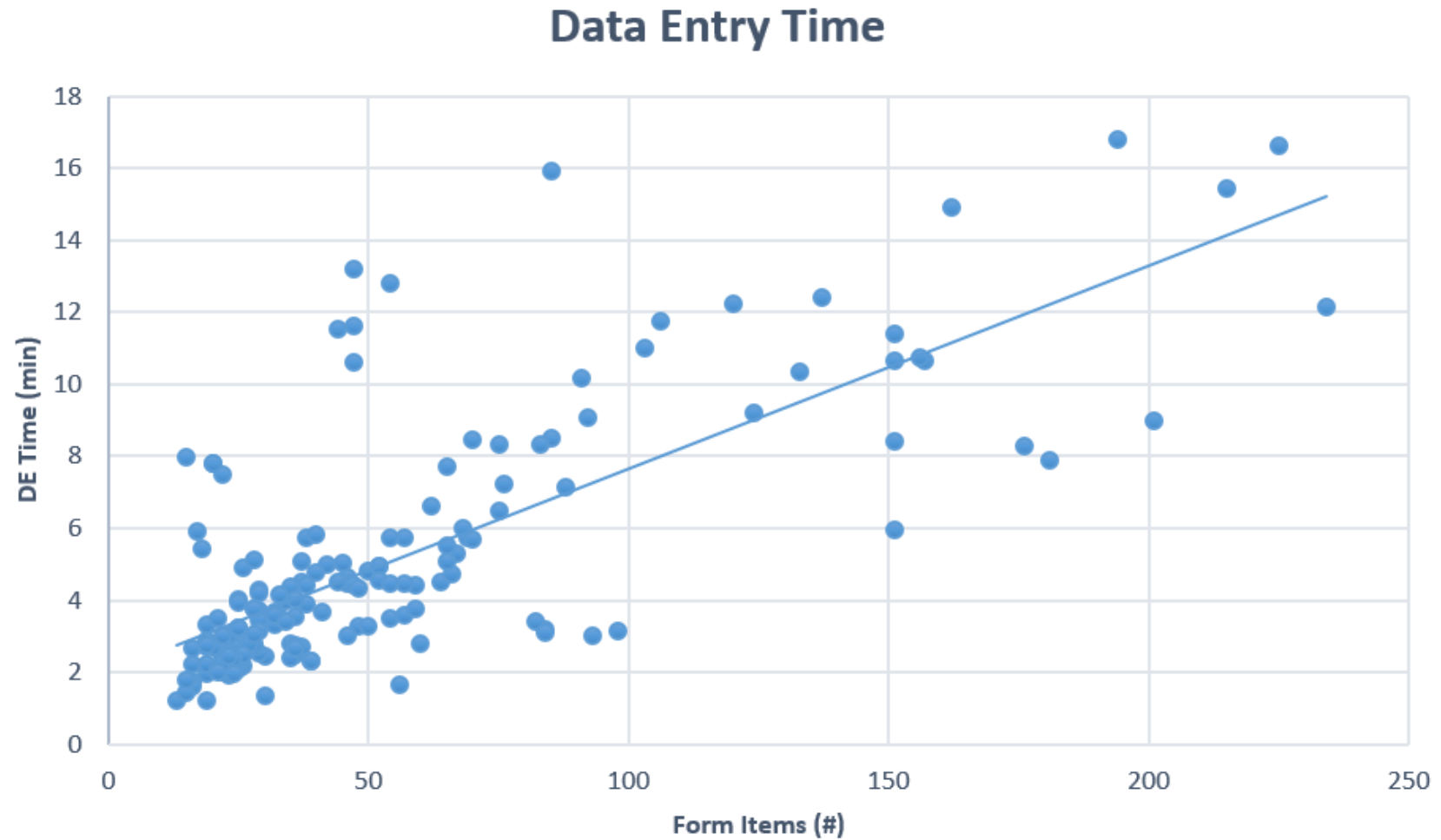
- Yes/No
- Single choice
- Multiple choice
- Numeric
- Date
- Text

Data entry time

- Double data entry
- Overall: 3.9 minutes/form
 - Alzheimer's studies: 4.3 minutes/form
 - Asthma studies: 3.9 minutes/form

	Min/form
Overall	3.9
Alzheimer's	4.3
Asthma	3.9

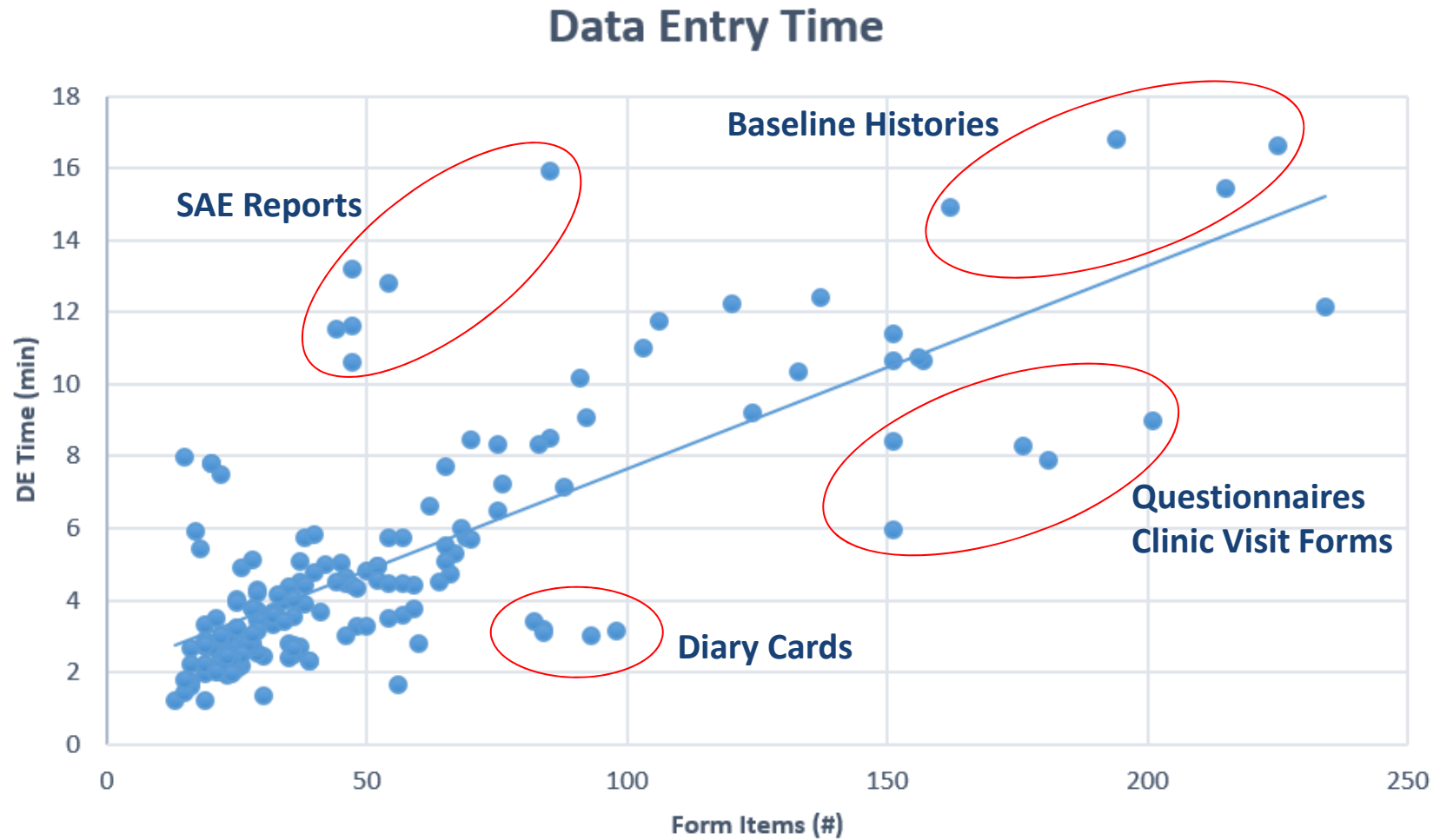
Data entry time vs # form items



Data entry time and item types

- In several analyses, there was little relationship between item type and data entry time
 - Other than multi-line open narrative blocks
 - Slight difference favoring non-character types
 - Confounded by multiple-choice items

Data entry time vs # form items



Data entry burden

- Overall burden of data entry averaged 342 minutes / randomized participant
 - Depends on number of visits and number of forms per visit
 - Likely a small fraction of the overall time spent / participant

Discussion

- Distribution of item types relatively consistent
 - Over time
 - Across disease-specific studies
- Relationship between data entry time and # of items relatively linear
- Little relationship between item type and data entry time
- Keying time subsequent to completion of paper CRF relatively low (3.9 min/form) even with DDE
 - Forms with long narratives take longer
 - Overall burden significant but only small part of effort