

第2回 NDB ユーザー会 (2019/8/23)
マスキング練習シート

奈良県立医科大学
野田 龍也

ガイドラインによる公表ルール (ガイドライン p28)

(1) 最小集計単位の原則 (注: 実際には①以外にもあります。)

- ① 公表される研究の成果物において、患者等の数が原則として10未満になる集計単位が含まれていないこと。

→要するに、「10未満」の患者数がわかる箇所をマスキングする必要があります。

実際の練習:

下記の表のうち、上記の①(10未満ルール)によりマスキングされるセルを指摘してください(例:A1)。

(A~E, 1~4は練習用であり、実際に審査へ出す際には不要です。)

(実際に審査へ出す際には数値が見えないようにする必要があります。)

(この表は創作されたダミーデータです。)

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18	0	1	19
B	1歳	12	0	14	26
C	2歳	30	4	3	37
D	3歳	38	31	50	119
E	合計	98	35	68	201

2019/8/23

【解説・解答】 マスキング練習シート

奈良県立医科大学

野田 龍也

1. 10 未満のためマスキング

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18	0	1	19
B	1歳	12	0	14	26
C	2歳	30	4	3	37
D	3歳	38	31	50	119
E	合計	98	35	68	201

2. 数値 0 も「10 未満」のためマスキング対象となる。
つまり、0～9 は問答無用でマスキング。

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18	0	1	19
B	1歳	12	0	14	26
C	2歳	30	4	3	37
D	3歳	38	31	50	119
E	合計	98	35	68	201

3. 「B1+B3」と B4（合計）の差分から B2=0 だと分かってしまう（一意に定まる）。
「逆算により 10 未満の数値が一意に定まる」場合もマスキング対象となる。
よって、B1,B3,B4 のいずれかをマスキングしなければならない（数値 10 以上でも）。
3 つのうち B1 が最小なので、B1 をマスキングしようか・・・

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18	0	1	19
B	1歳	12	0	14	26
C	2歳	30	4	3	37
D	3歳	38	31	50	119
E	合計	98	35	68	201

4. ところが、B1 をマスキングすると、「A1+C1+D1」と E1 の差分から B1=12 と分かってしまう。
B1=12 が分かると前項より B2=0 も導かれてしまう。
B1 ではなく、B4 をマスキングした場合も同様に B4=26, B2=0 が逆算できてしまう。
となると、B 行は B3 をマスキングすべきである。

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18	0	1	19
B	1歳	12	0	14	26
C	2歳	30	4	3	37
D	3歳	38	31	50	119
E	合計	98	35	68	201

5. マスキング終了。

合計項目があるタテヨコ（行・列）では、1ヶ所だけマスキングがある場合は逆算可能となるので、要注意。

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18			19
B	1歳	12			26
C	2歳	30			37
D	3歳	38	31	50	119
E	合計	98	35	68	201

6. せっかくの10以上を消すのはもったいない。

B3をマスキングするのではなく、「15程度」という幅のある表現とすれば、B2=0とはわからなくなる。

野田流マスキング終了。

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18			19
B	1歳	12		15程度	26
C	2歳	30			37
D	3歳	38	31	50	119
E	合計	98	35	68	201

7. 注意！

Excel で「文字色とセルの塗りつぶしを同色」にしても、印刷すると読めてしまいます。マスキング前の数値を残したい場合は、字色も塗りつぶしも真っ白または真っ黒にします。

なお、PDF はテキスト選択不可処置を行う必要があります。

(下は文字色とセルの塗りつぶしを灰色にした例)

		1	2	3	4
		Pがん	Qがん	Rがん	合計
A	0歳	18			19
B	1歳	12			26
C	2歳	30			37
D	3歳	38	31	50	119
E	合計	98	35	68	201

「10 未満マスキング」のコツのまとめ

- ・ 0～9 の患者数はマスキングが必要。0（不存在）もマスキングが必要。
- ・ 逆算して「0～9 の患者数」が「一意」に決定できる場合もマスキングが必要。
→0～9 の間の「特定の数値」であることが推定できる場合はアウト
→「数値を 1 通りに特定できないが、10 未満であることが分かる」場合はセーフ。
→「程度」といった幅のある表現を用いて、基準を守りつつ情報量をなるべく守る。
- ・ 「合計」「小計」がある行・列でマスキングが 1 ヶ所だけある場合は、10 未満を逆算可能となるので、要注意。
- ・ 遠隔したセルや複数表の組み合わせで合計が分かる場合にも、上記と同様の注意が必要。(例：男性、女性、合計で表が分かっている場合)
- ・ Kaplan-Meier 曲線の「段差」で 10 未満の特定の数値であることが分かってしまう場合もマスキングが必要であるなど、意外なところで基準に引っかかることはある。