T1International

Survey 2022 - Data cleaning & initial silo analyses

Axel Thieffry

November 2022

Content

- Simplification of question names

- Data cleaning: cumulative filtering

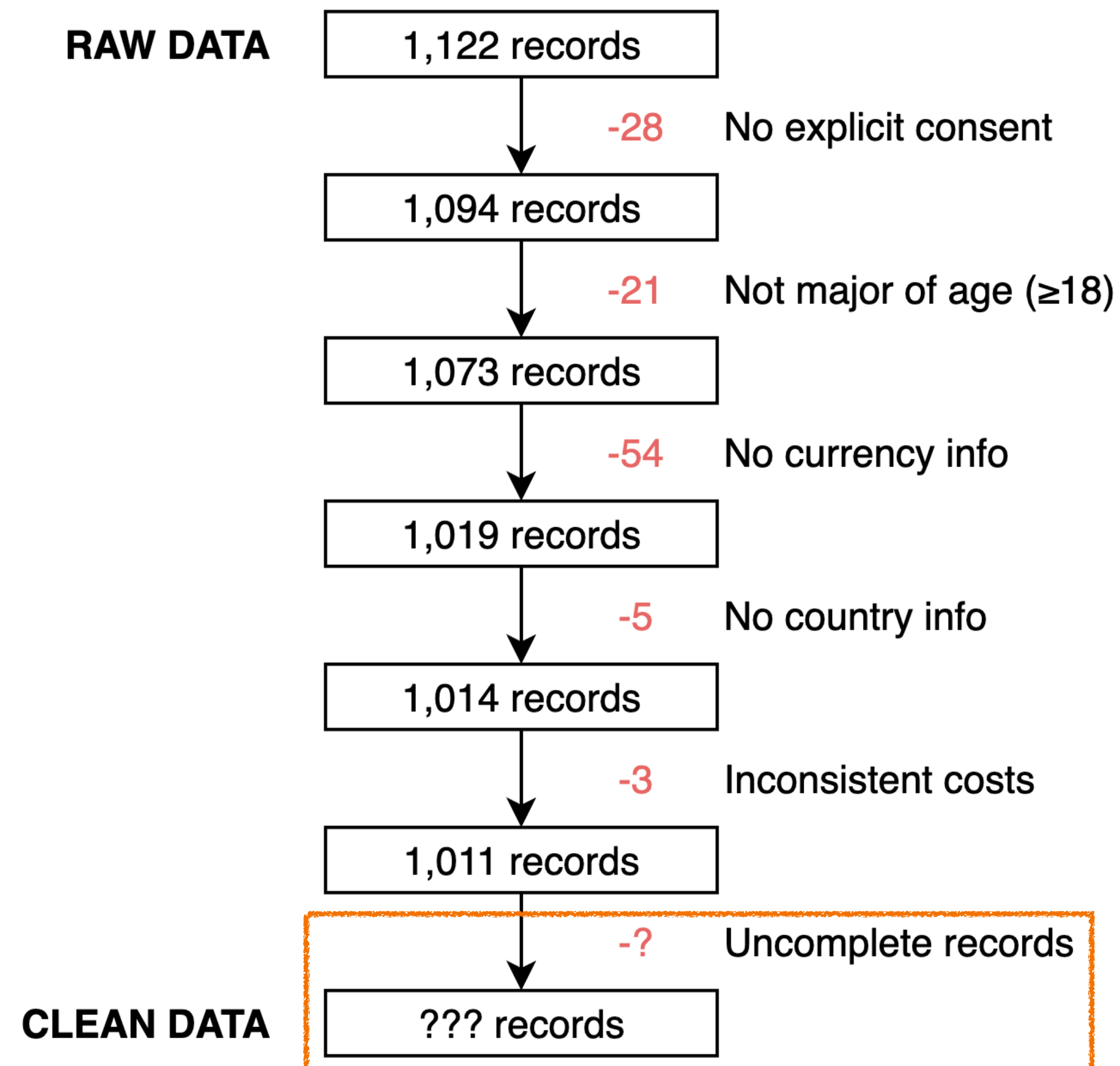- Standardise costs to USD

- Uncomplete records

- Silo analyses

Simplification of question names
for easier data handling

- Available in a spreadsheet

- Will appear in draft figures

- Will be polished in final figures

renamed_column	initial_name_column
help_no	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=No)
help_famfriends	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Yes, su
help_charity	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Yes, ch
help_donation	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Yes, do
help_gov	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Yes, go
help_pharma	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Yes, ph
help_other	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Other {
help_pnta	Do you receive any help to pay for your insulin, diabetes supplies, or care? (Choose all that apply) (choice=Prefer
help_other_desc	If other, please describe
payOoPE_none	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_savings	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_borrow	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_asset	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_other	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_pnta	In the past year, have you had to do any of the following to pay for your out-of-pocket costs for your medication an
payOoPE_other_desc	If other, please describe
covid_unaffected	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=No change)
covid_supply	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Access to supplies ha
covid_insulin	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Access to insulin has
covid_price_up	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Prices have gone up)
covid_price_down	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Prices have gone dow
covid_unsure	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Unsure)
covid_pnta	Has the COVID-19 pandemic affected your access to insulin and diabetes supplies? (choice=Prefer not to answer)

Data filtering

RAW DATA

1,122 records

-28   No explicit consent

1,094 records

-21   Not major of age (≥18)

1,073 records

-54   No currency info

1,019 records

-5   No country info

1,014 records

-3   Inconsistent costs

1,011 records

-?   Uncomplete records

CLEAN DATA | ??? records |

- Cumulative filtering (top to bottom): not absolute numbers as some entires might have been already discarded from previous step(s)

- Uncomplete records to remove: TBD

Standardise costs to USD
a. currency extraction

- Currency symbol (ISO 4217, ex. "(GBP)") extraction strategy:

    1. If declared from drop-down menu: directly from answer

    2. If other currency ("Your Currency"): taken from description after manual curation

- Manual curation:

    - New column in raw data (see example below), for 112 cases

    - Country info used to resolve ties

Record ID	What country do you live in?	What currency do you use?	If your currency, please specify	currency_other_desc_manual
130	Argentina	Your Currency	Peso Argentino	ARS
966	Colombia	Your Currency	Peso coombiano	COP
1032	Colombia	Your Currency	pesos colombiano	COP
1002	Costa Rica	Your Currency	Colones	CRC
461	Denmark	Your Currency	Dk	DKK

Standardise costs to USD
b. currency conversion

- Conversion to USD uses current FX rates from the World Bank API (via the priceR package)

- All initial cost-related columns with format: `cost_xxx`

- All USD-converted columns renamed with format: `usd_xxx`

Standardise costs to USD

c. sanity check

- Manual sum of Out-of-Pocket Expenses (OoPEs)

  vs

  Survey's automatic calculation

- Definition:

$$OoPE_{total} = \sum (insulin_{[sa+la+mix+other]}, pump, bg, cgm, pen \& needles)$$

⇒ Survey automatic calculation is good

⇒ 3 outliers removed

Costs sanity check
Manual cost sum vs survey−calculated cost sum
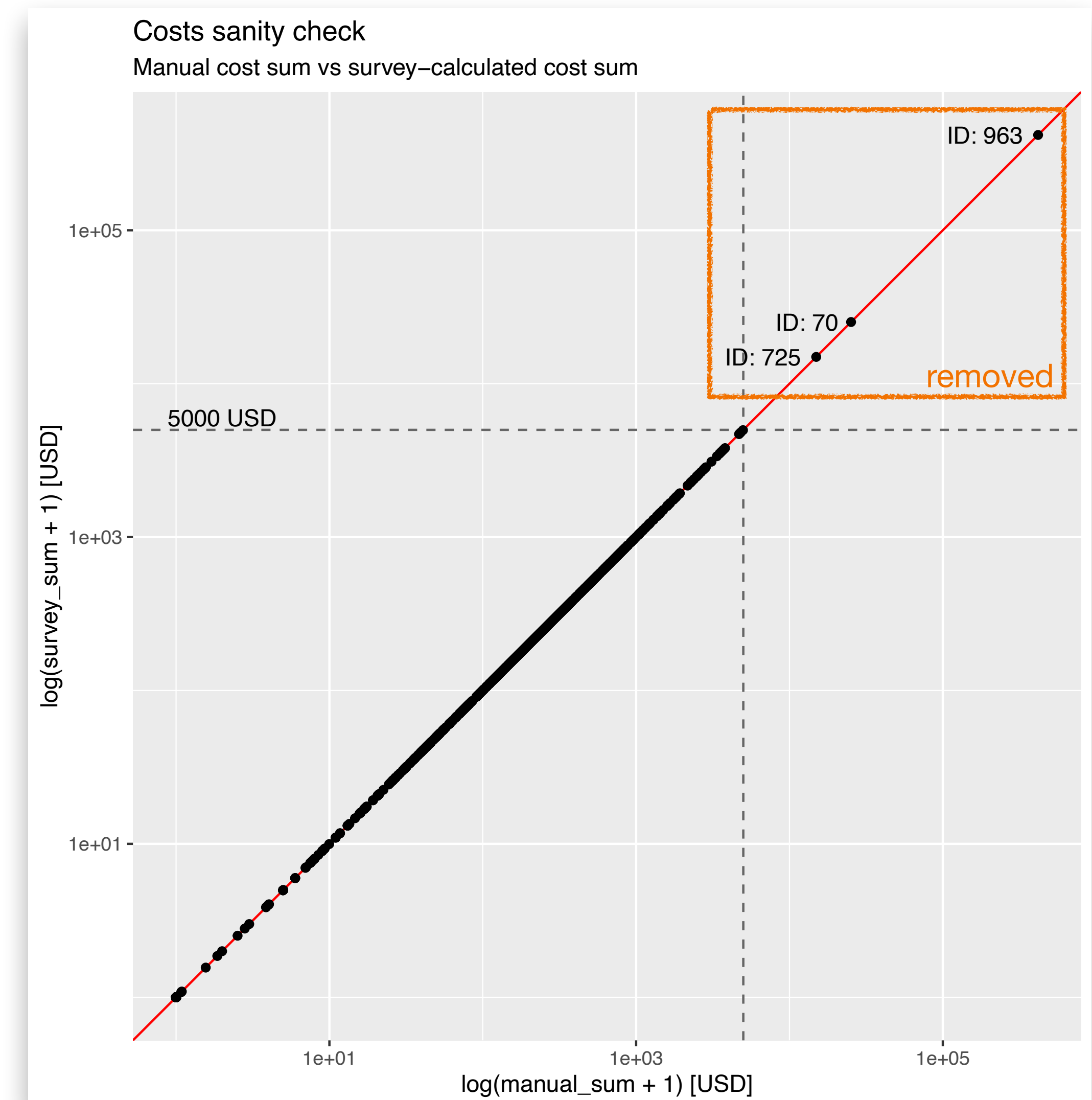
Uncomplete records
Overview

- Currently **27** records are marked as *incomplete*

- Tentative set of important* questions:
  - T1 connection, country, healthcare coverage
  - rationing of insulin and BG testing
  - all T1D-related costs (OoPEs)
  - other cost-related quesions (doctor, premium, household)

⇒ Keeping the 27 *incomplete* records won't have dramatic consequences as many of the *completed* records also failed to inform on many of these questions
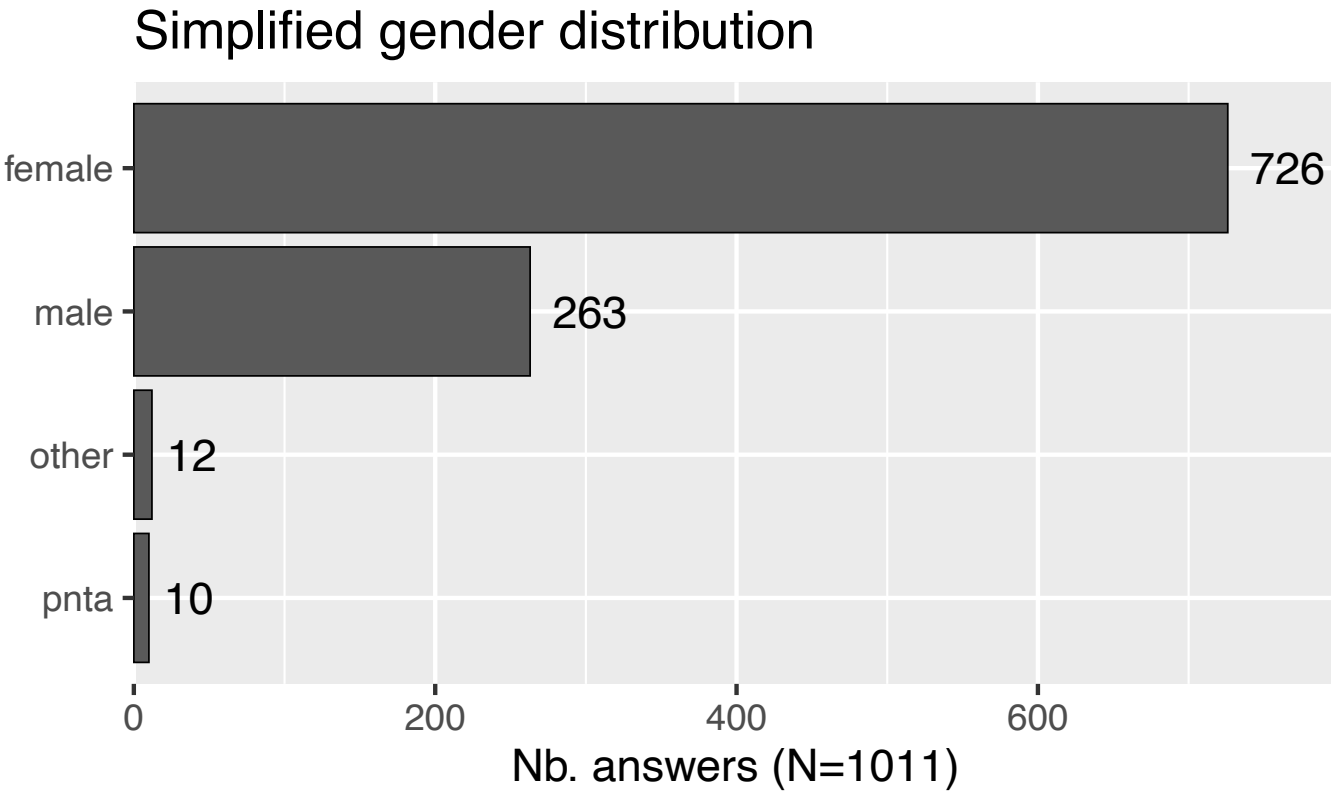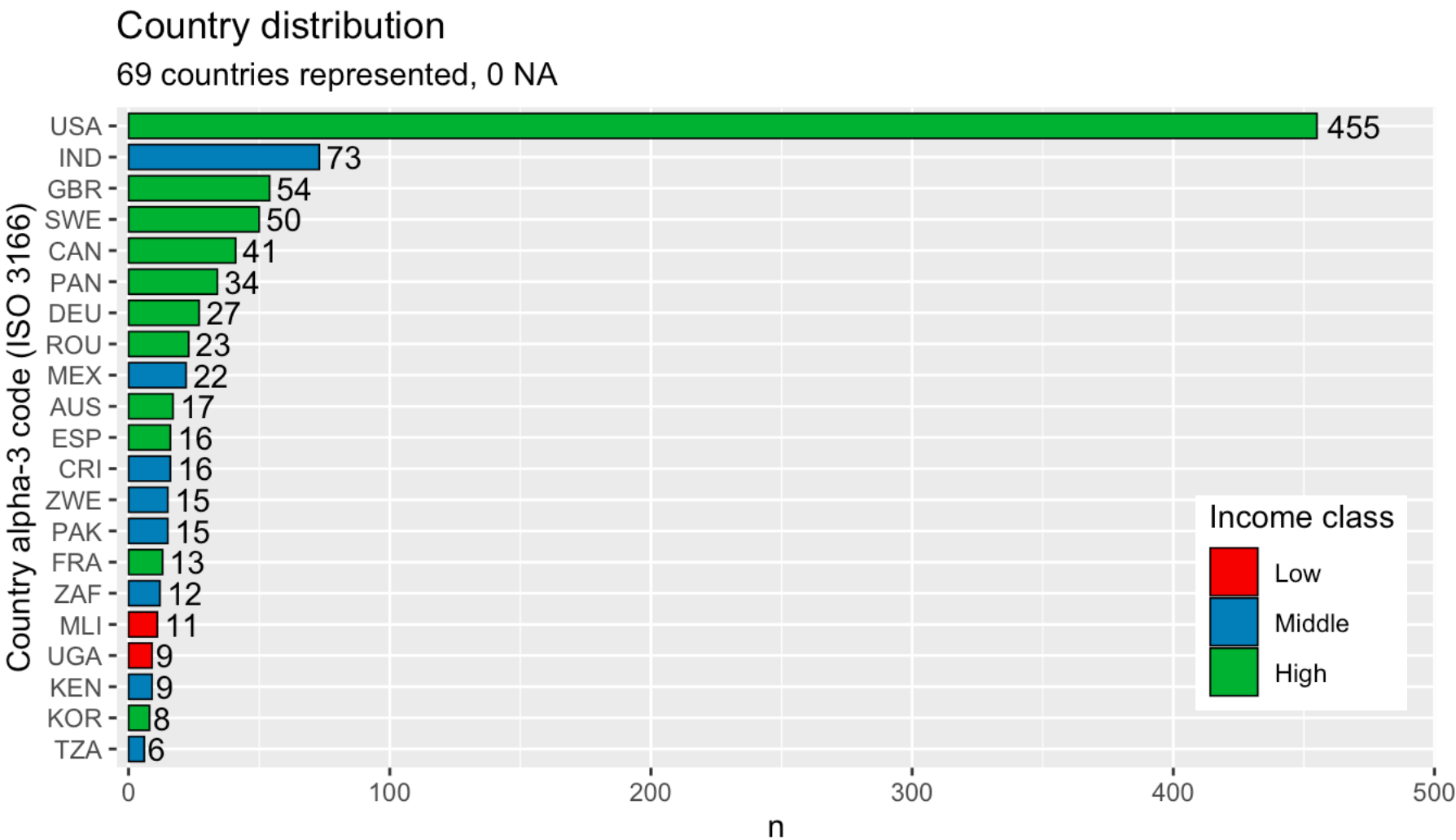
⇒ If kept, final dataset will comprise **1,011** answers

arbitrarily-defined as per the intended main goals of the survey
e.g. costs, COVID, and rationing
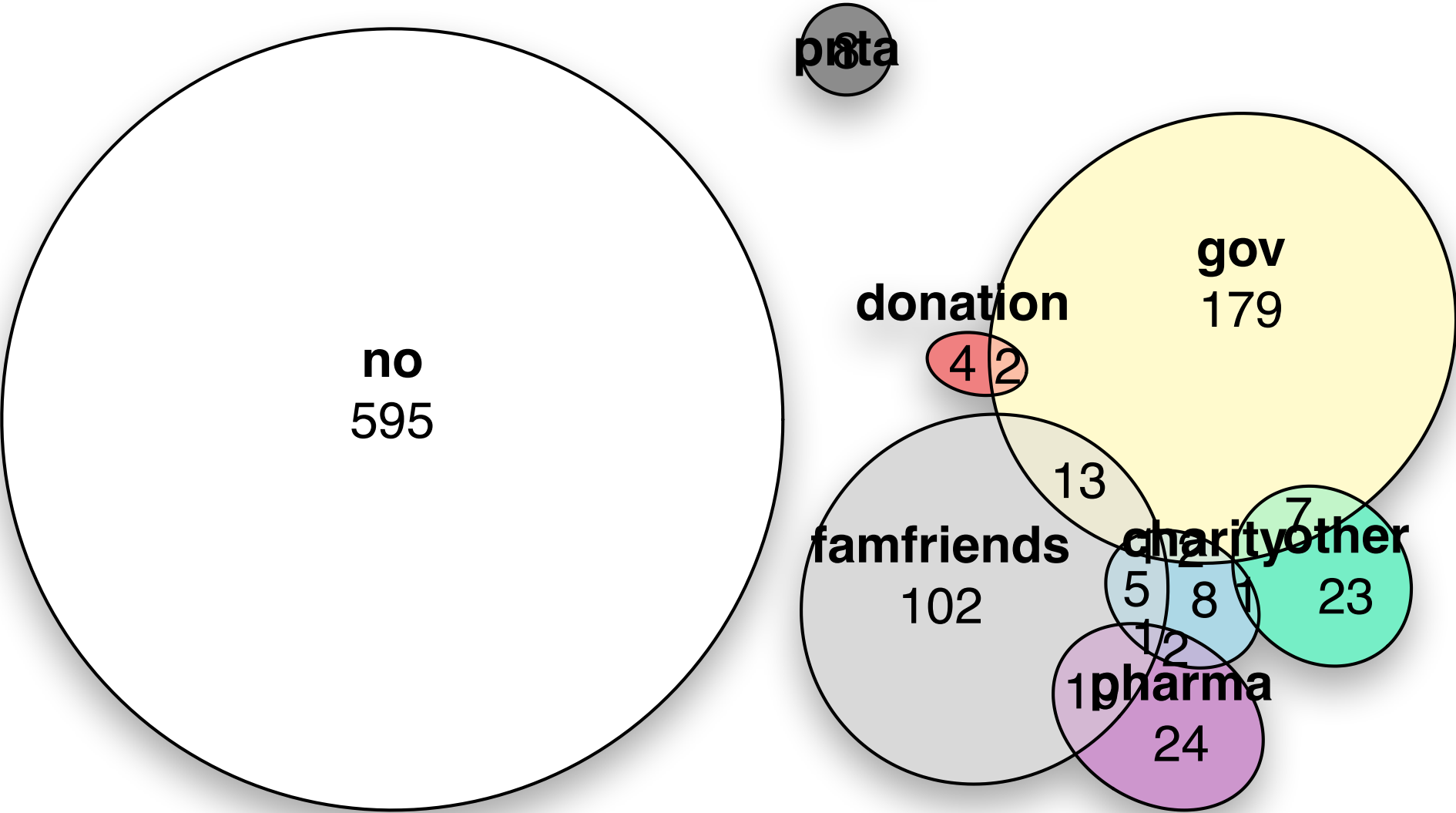

8
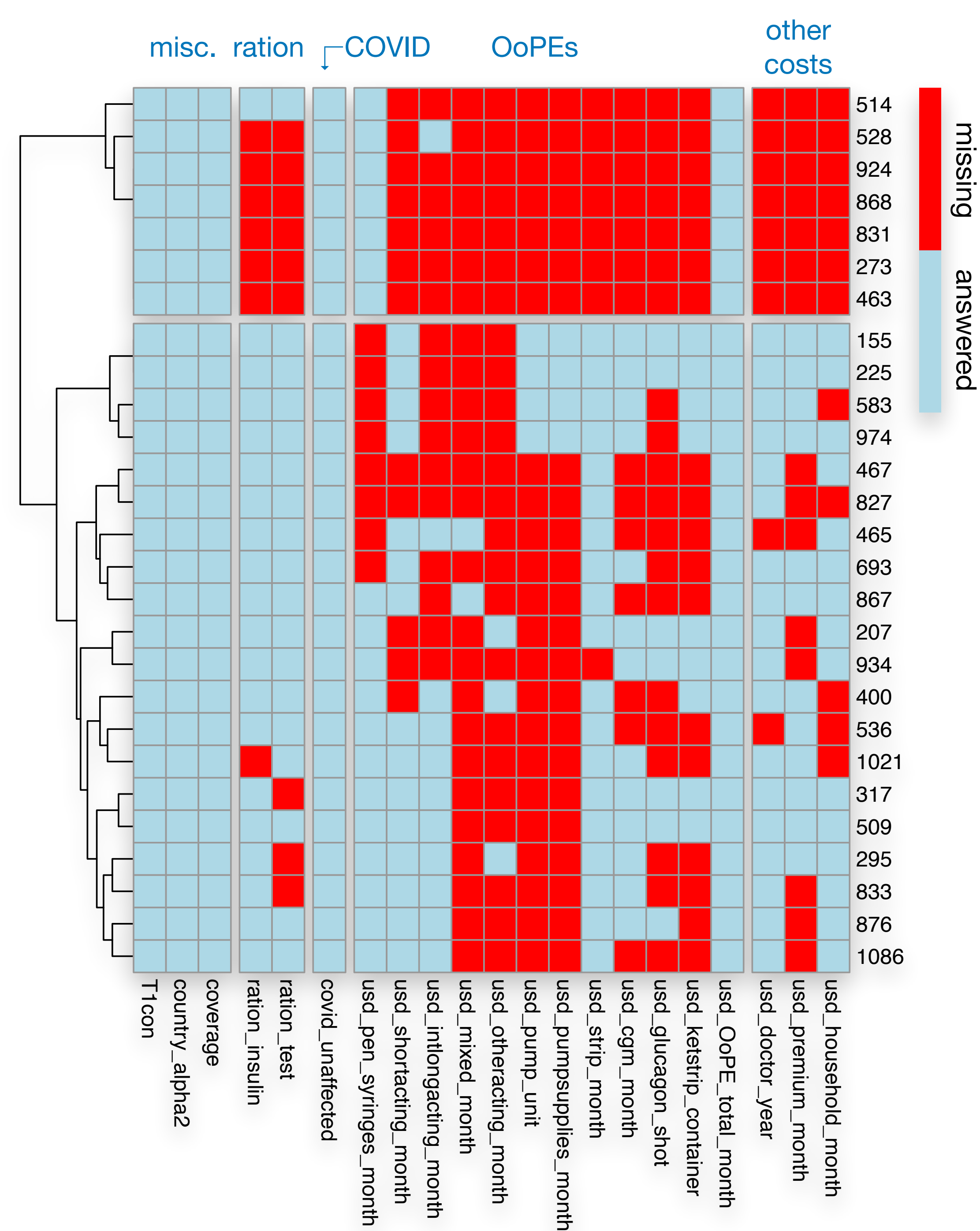
Silo analyses
Demographics & misc

Country distribution
69 countries represented, 0 NA

Simplified gender distribution

Category	Count
female	726
male	263
other	12
pnta	10

Nb. answers (N=1011)

Connection to T1D

Category	Count
patient	768
mychild	224
betterhalf	13
doc	4
pnta	2

Nb. answers (N=1011)

Healthcare coverage distribution

Category	Count
partial	669
none	177
full	160
pnta	5

Number of answers (N=1011)

Race

white 629
mideast 17
asian 103
black 73
indig
hisp 102
other
pnta 19

⚠️ *Note:*
Euler diagrams (cfr. above) are comparable to classical Venn diagrams except:
- Areas are proportional to the number of members in a set
- Must be interpreted as an overview rather than an exact representation of the data reality: all set intersection are not always possible to represent geometrically
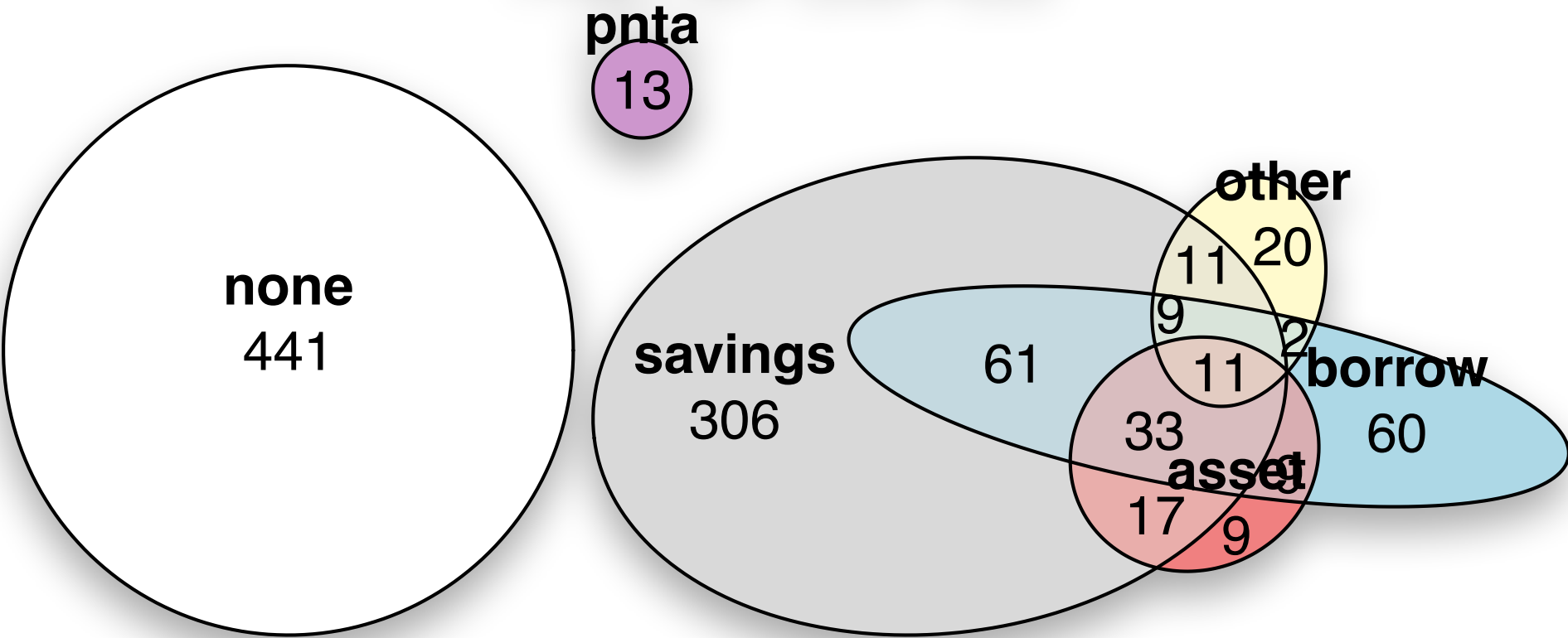
Silo analyses
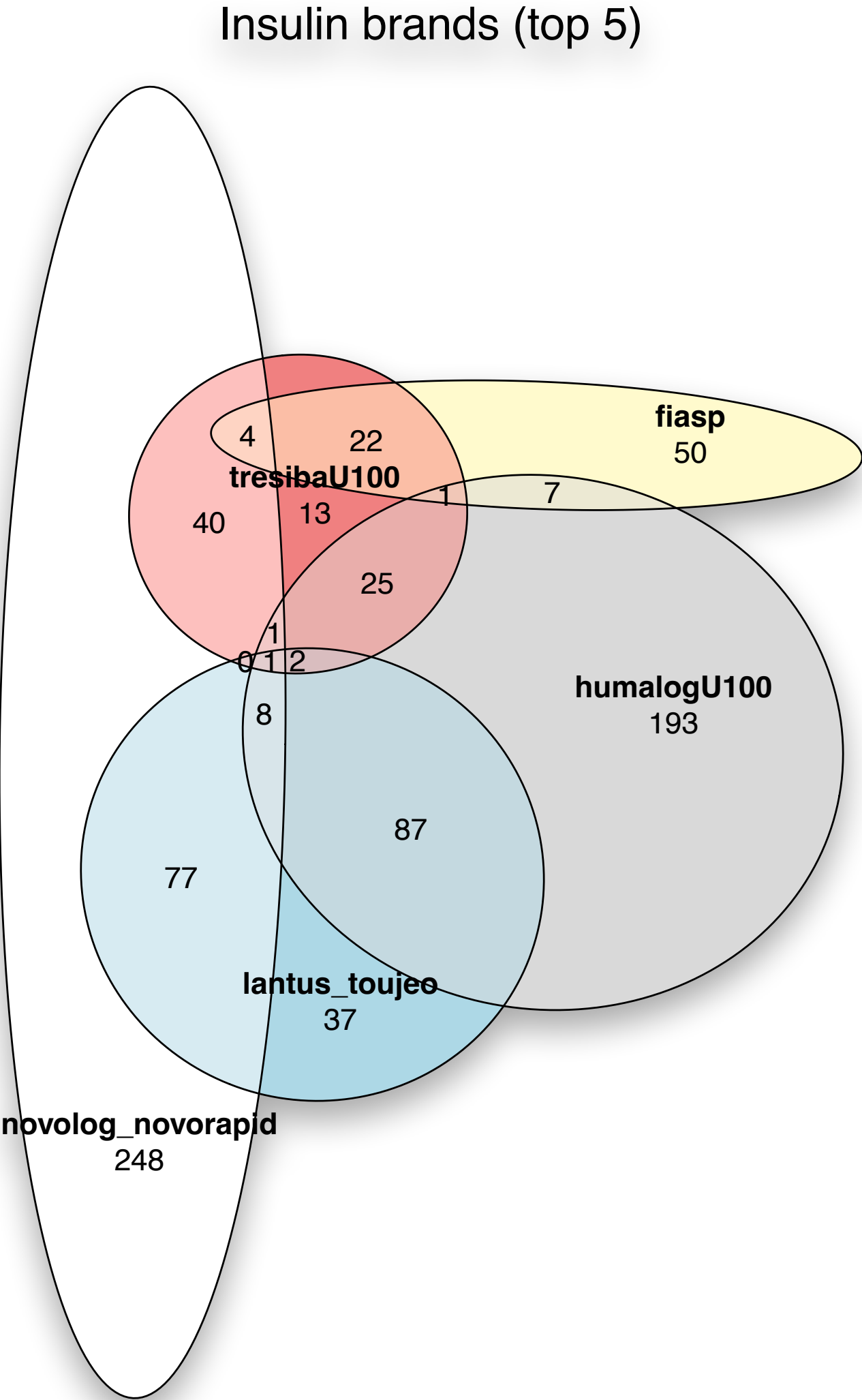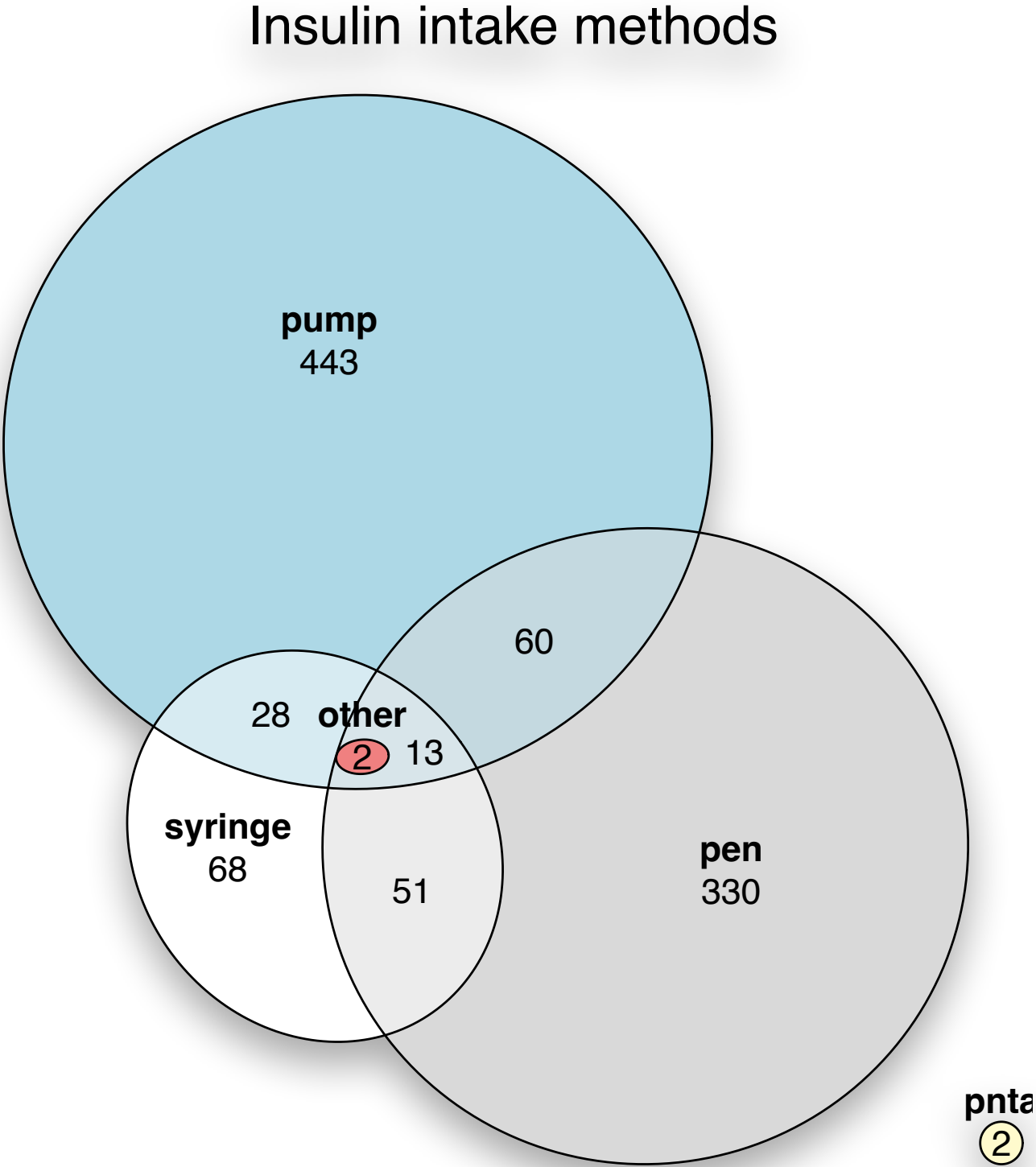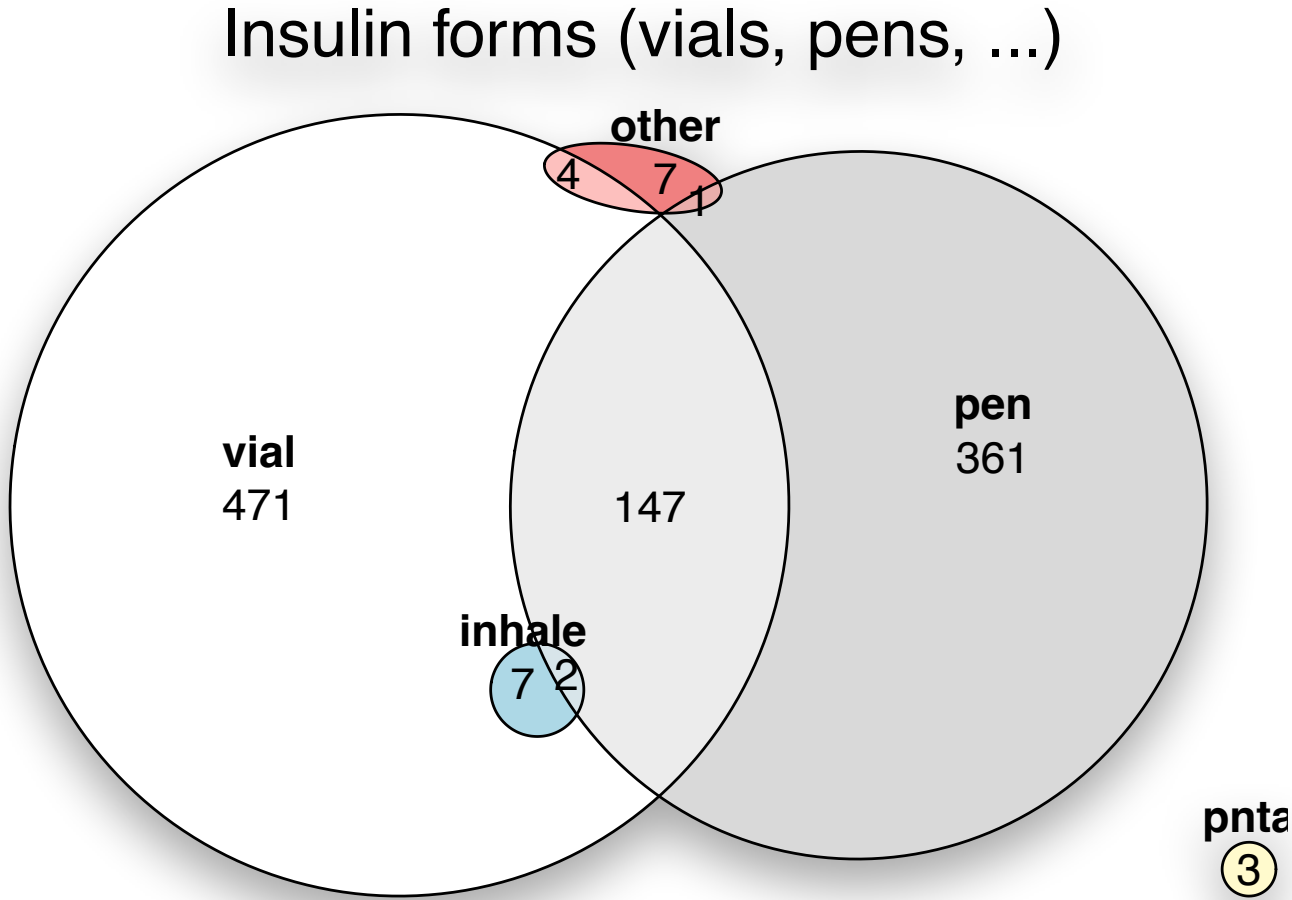## Financing

Financial help

pnta 8

no
595

donation
4 2

gov
179

13

famfriends
102

charity
5

other
8 7

7

other
23

1 2

1 pharma
24

Pay for OoPEs

pnta
13

none
441

savings
306

other
11 20

9

61

2

11

borrow
60

33

asset
17 9

Silo analyses

Insulin

Insulin brands (top 5)

Insulin intake methods

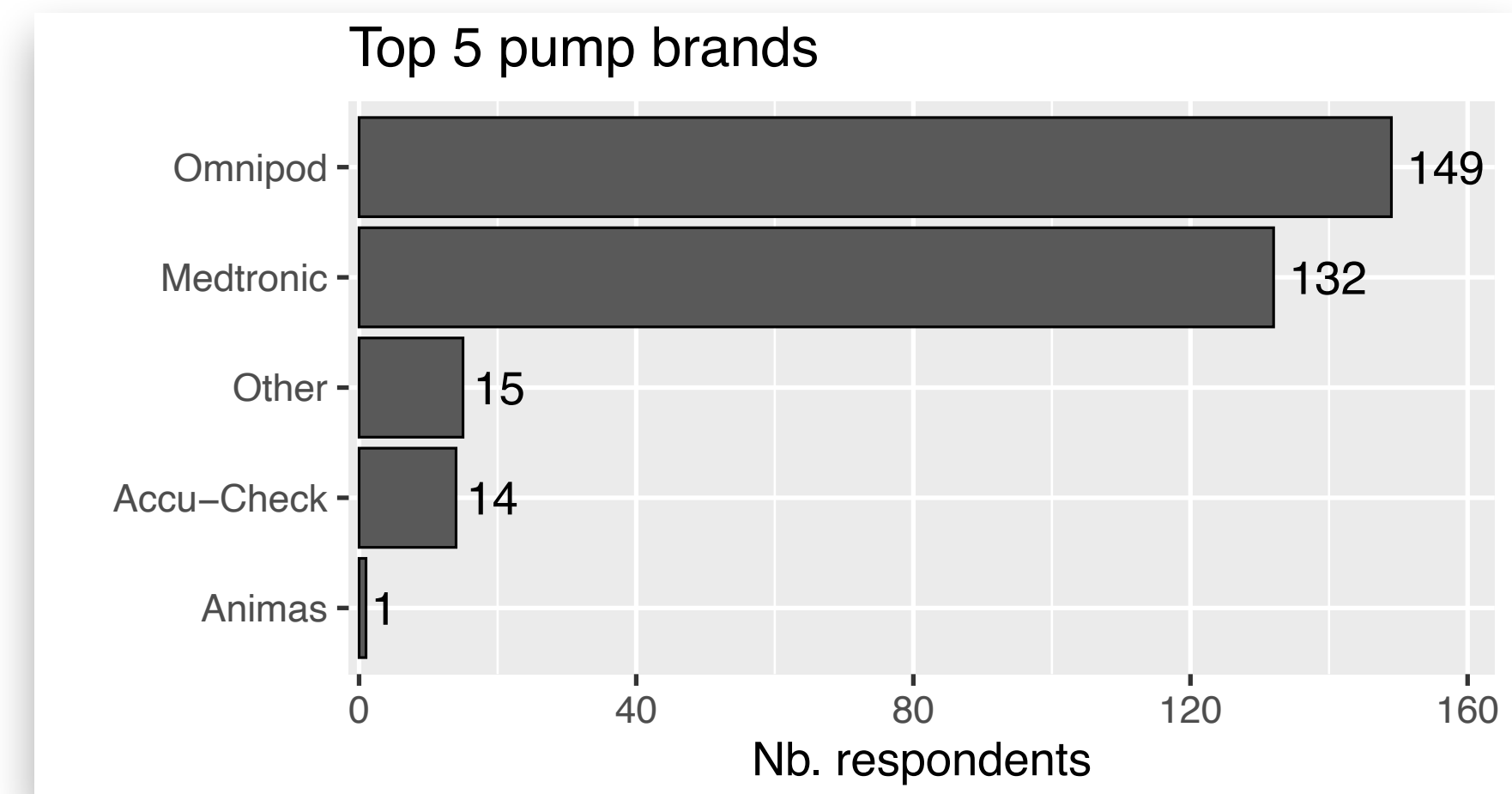Insulin forms (vials, pens, ...)

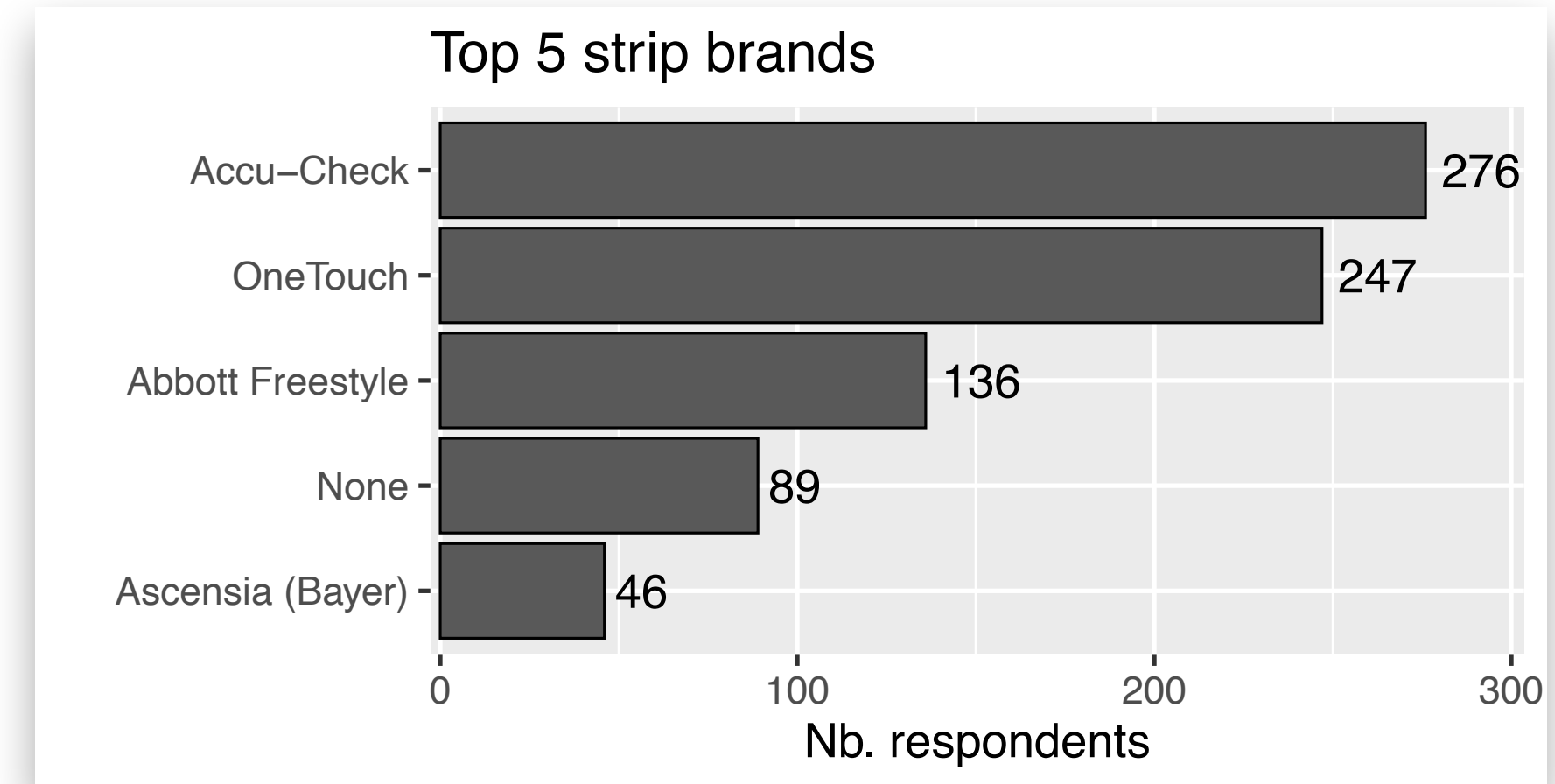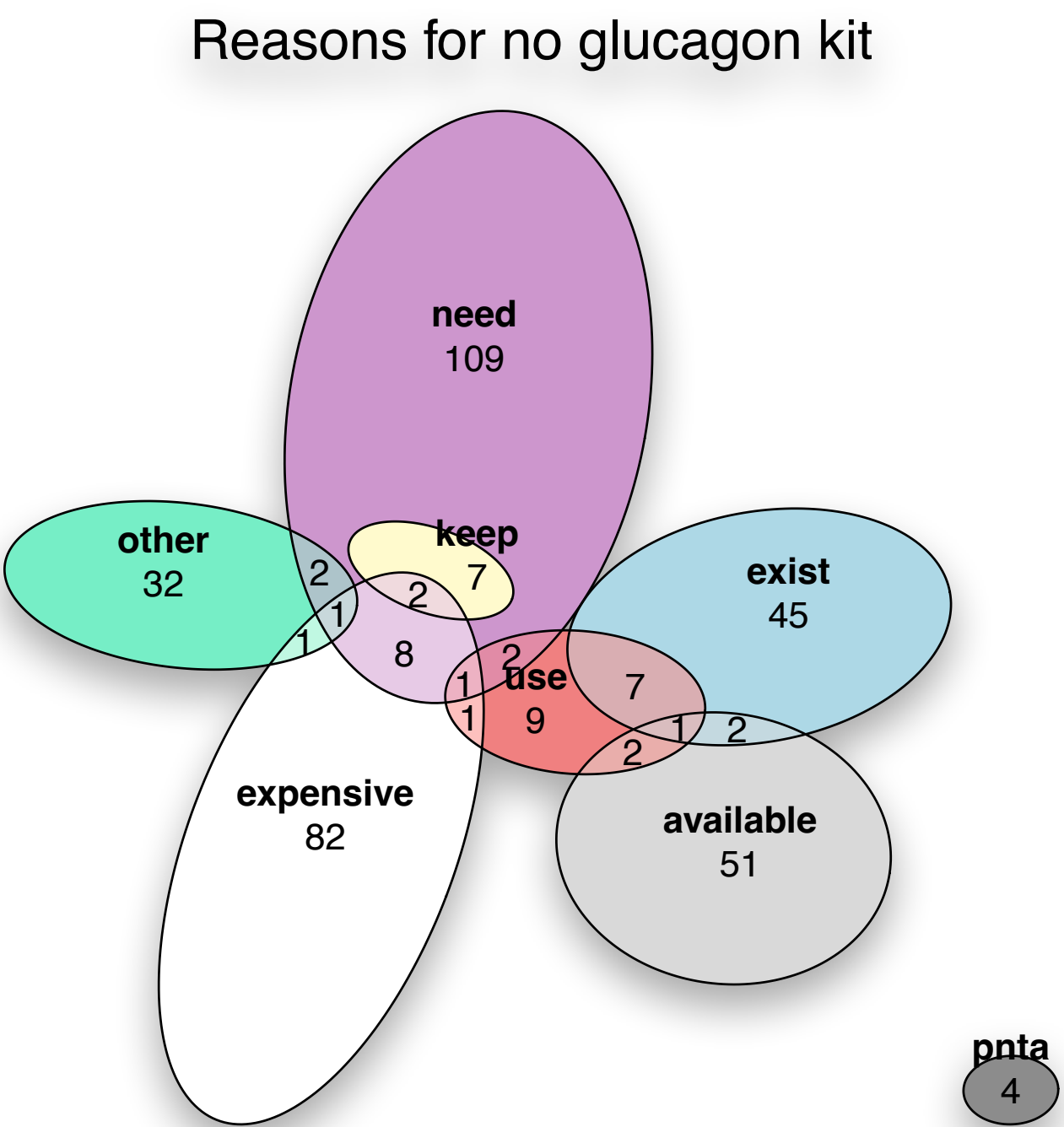# Silo analyses
Rationing
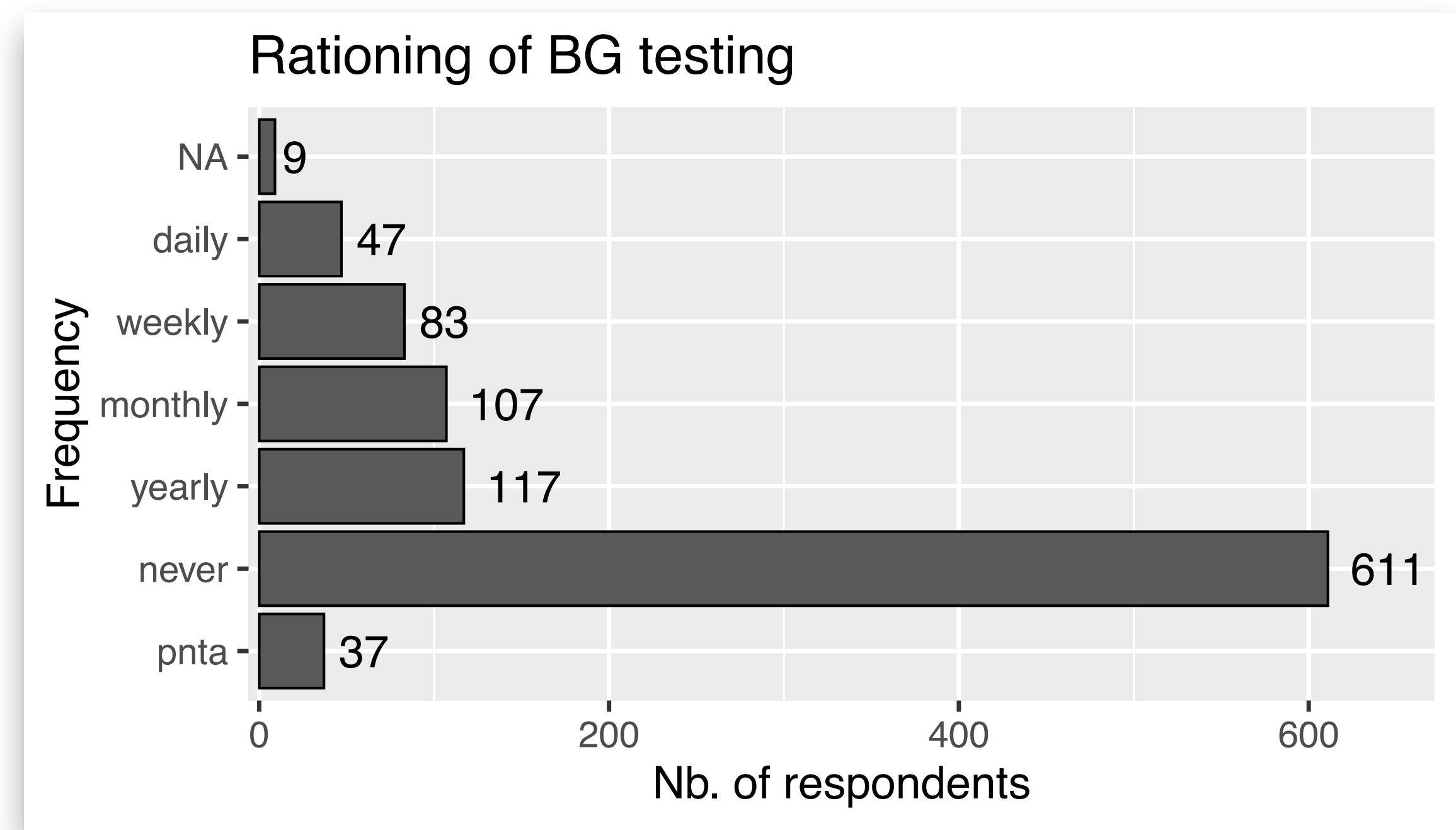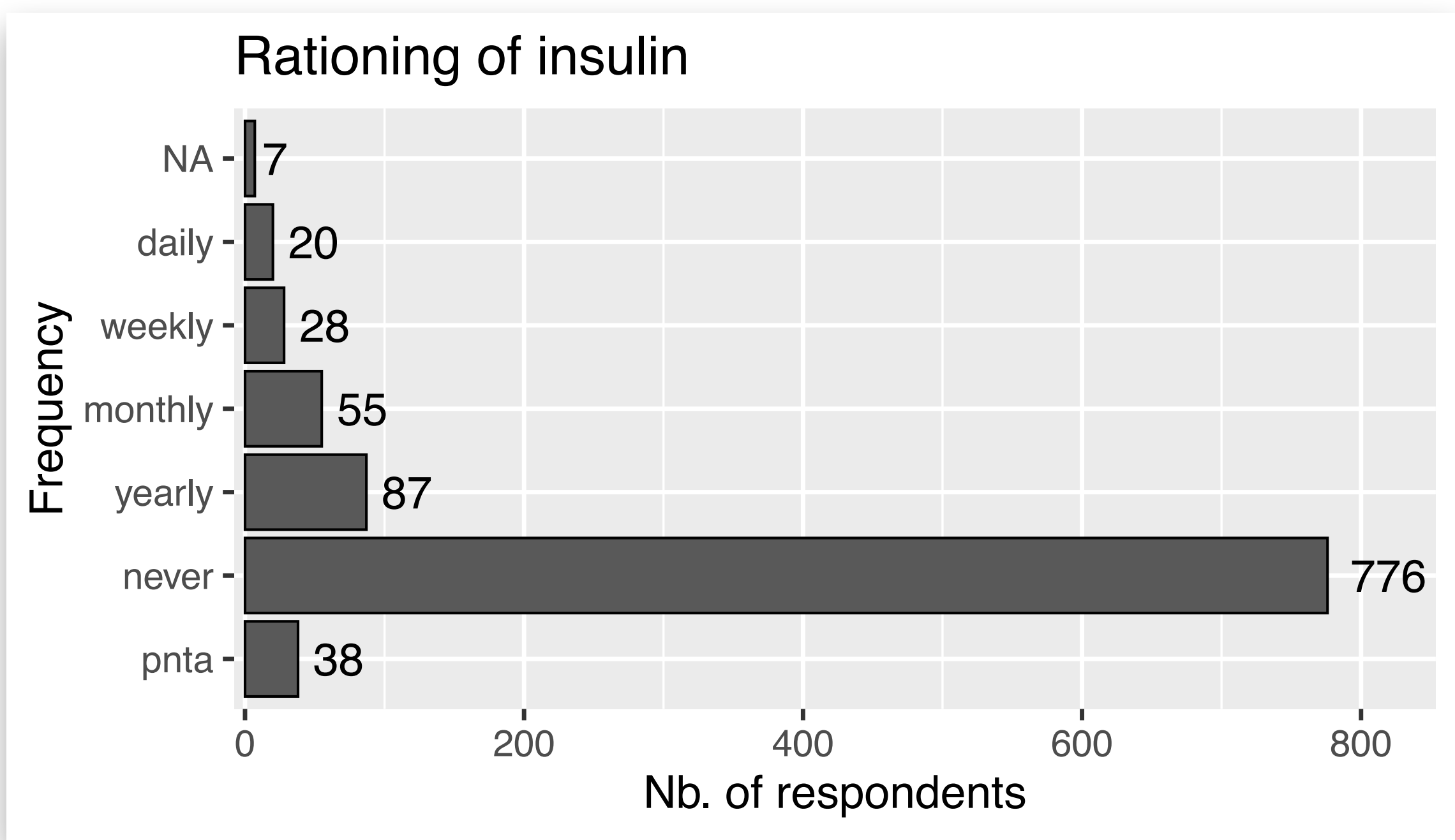
Silo analyses
COVID-19 impact
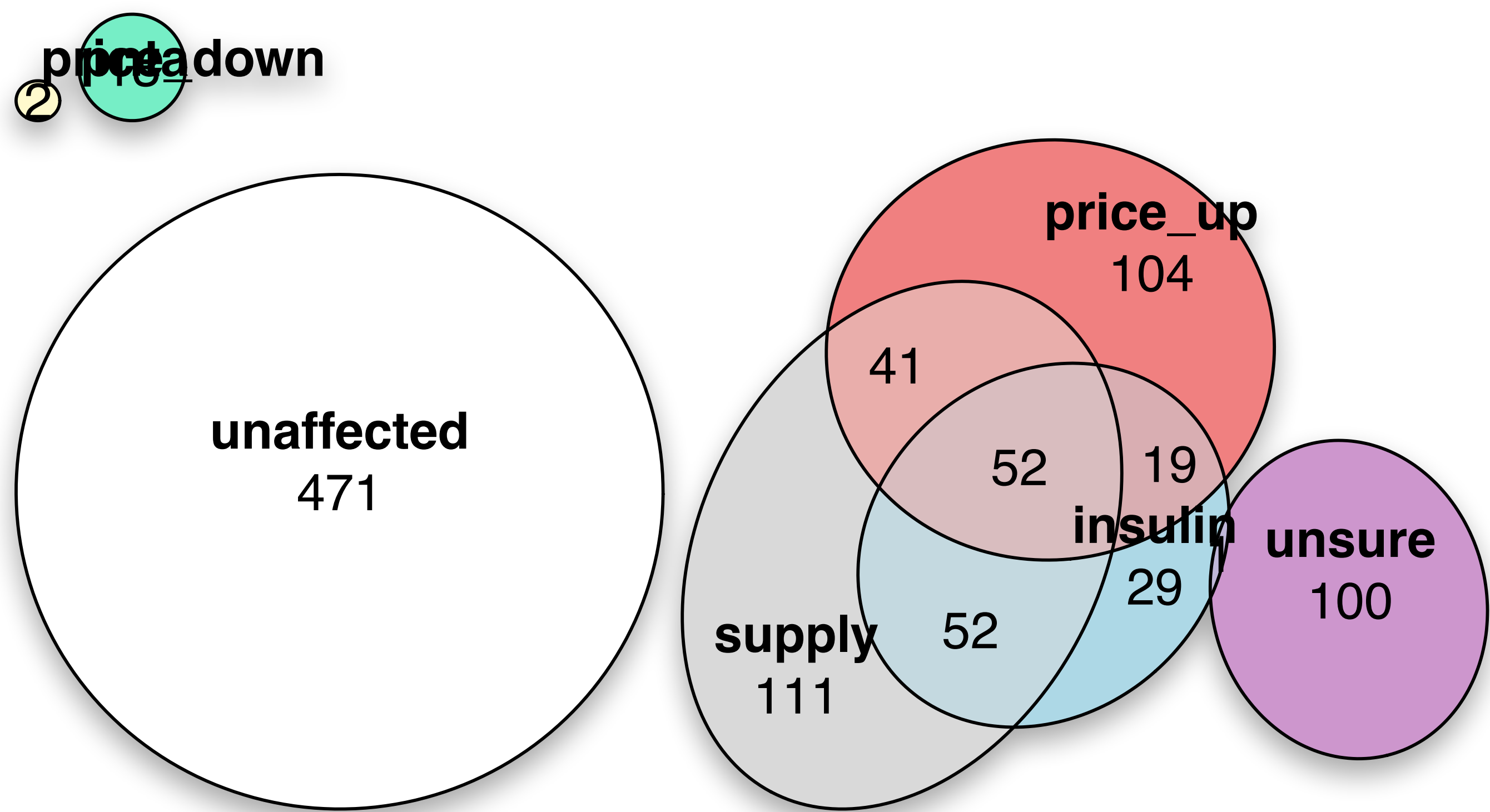
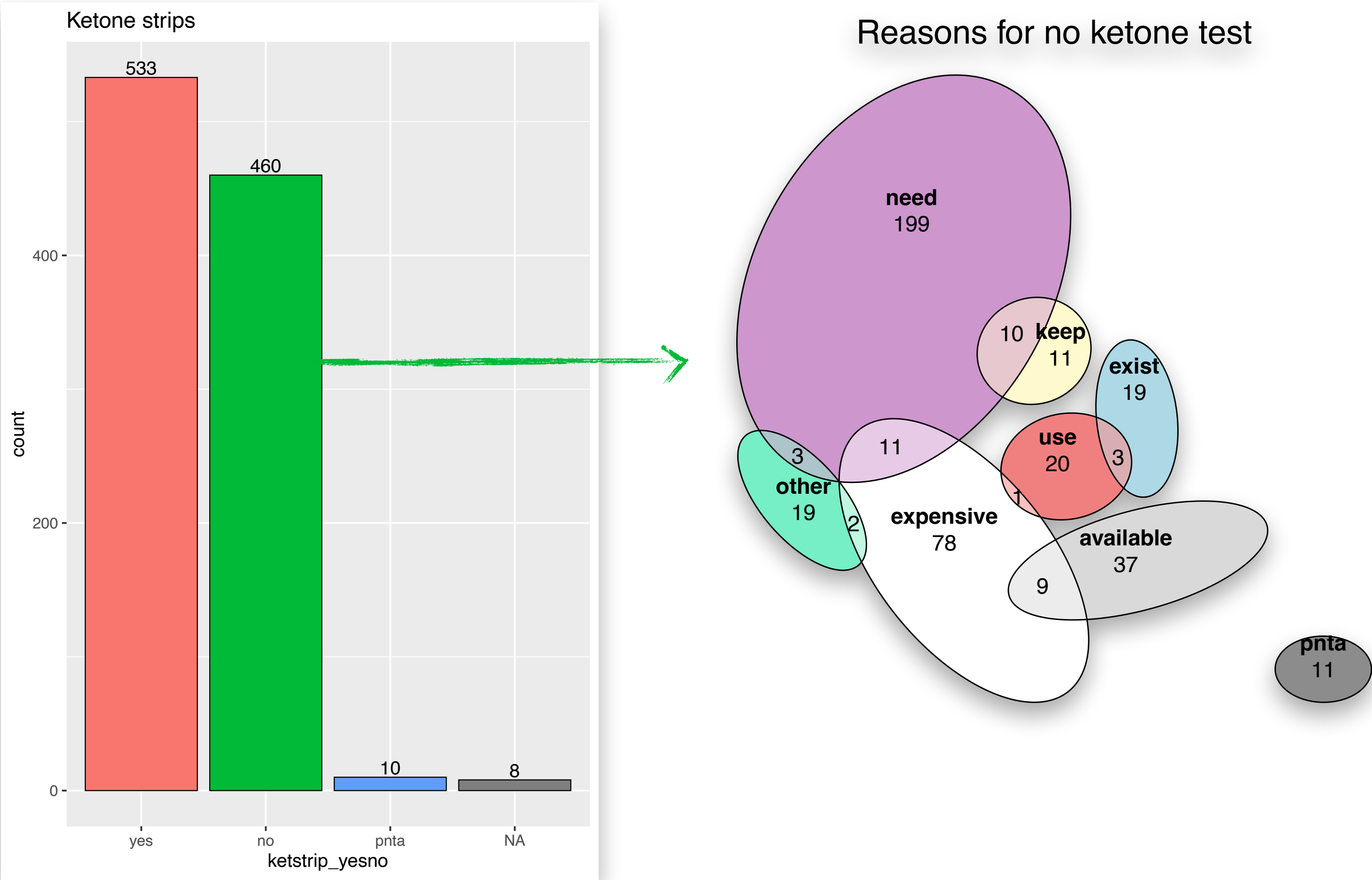# Silo analyses
CGM, test strip, and pumps

CGM usage and brands

Top 5 strip brands

Top 5 pump brands
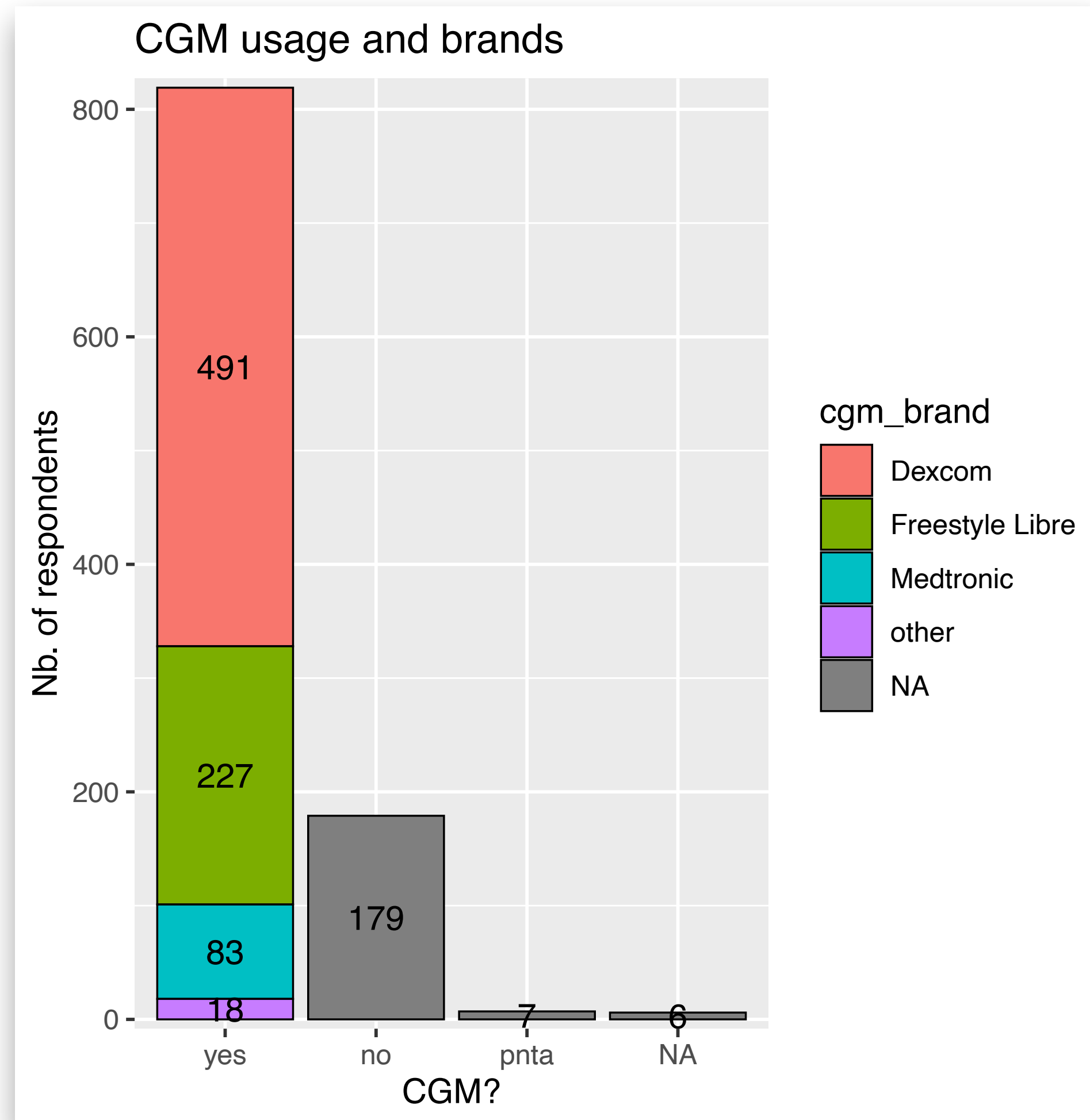
Silo analyses
Glucagon kit

Silo analyses
Ketone kit

Silo analyses
Cost distributions

Note:
0 USD are not shown because lost by the log transformation: log(0) = -∞

Silo analyses
Cost distributions

Note:
0 USD are not shown because lost by the
log transformation: log(0) = -∞
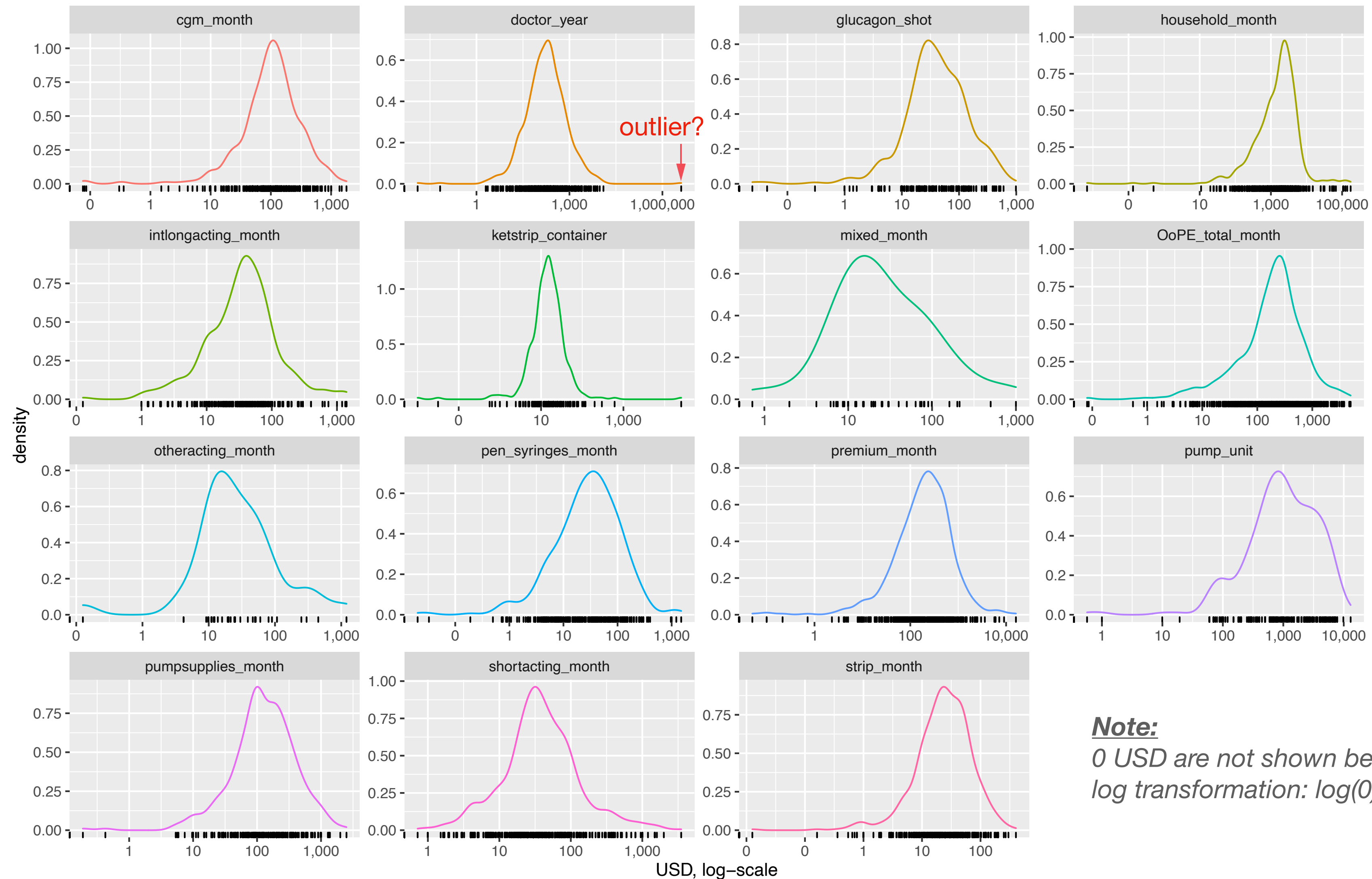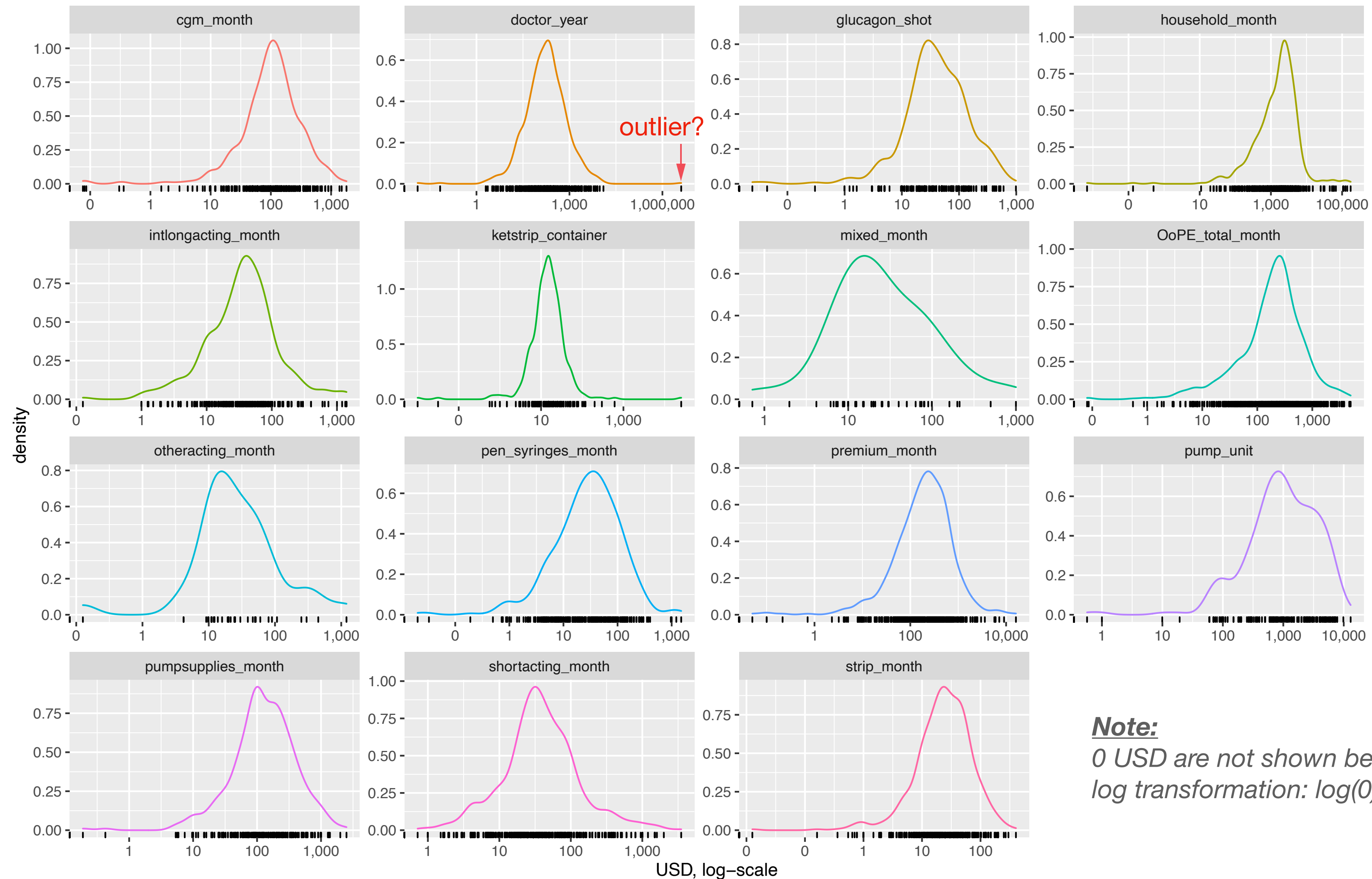
Survey inconsistencies: OoPEs

▸ Cases where a respondent indicated costs for an insulin-management method they say they don't use

▸ Found in:
   - Insulin pen & needle
   - Insulin pump
   $\Rightarrow$ See next slides

▸ Not found in:
   - BG test strips
   - CGM
   - Glucagon
   - Ketone strips

▸ Only two respondents: ID 807 & ID 903

▸ Action:
   a) costs were removed for these events
   b) total OoPE/month were recalculated for these two respondents

Survey inconsistencies: OoPEs

▸ 1: Usage & OoPEs for insulin pen & needles:

22	[insulin_how]	How do you administer the insulin? (Choose all that apply)	checkbox, Required		
			0	insulin_how___0	Syringes
			1	insulin_how___1	Pens with pen needles
			2	insulin_how___2	Insulin pump
			3	insulin_how___3	Other {insulin_how_oth}
			4	insulin_how___4	Prefer not to answer

24	[pen_needles_cost] Show the field ONLY if: [insulin_how(0)] = '1' or [insulin_how(1)] = '1'	How much do you pay in [currency] for a 1 month supply of pen needles and/or syringes?	text (number) Custom alignment: LV

```
 id insulin_intake_syringe insulin_intake_pen usd_pen_syringes_month
<dbl> <lgl>                <lgl>                              <dbl>
 807 FALSE                FALSE                                 30
 903 FALSE                FALSE                                121.
```

▸ Question of *cost for pen & needles (#24)* should only appear in the survey if respondents had indicated using *syringes or pens with pen needles* in question #22.

▸ But two respondents could indicate costs in question #24 without having the necessary 'checks' in question #22.

▸ *Note: perhaps the fact that some respondents not being 18+ age still had access to the whole survey originates from the same survey logic problem.*

▸ I see in *#24* that the values required for showing the field are named "insulin_how(0)" and "insulin_how(1)", whereas those values in *question #22* are named "insulin_how___0" and "insulin_how___1".
Could that be the reason why the survey logic did not work? If so, that problem is systemic.

Survey inconsistencies: OoPEs

▸ 2: Usage & OoPEs for insulin pump:

22	[insulin_how]	How do you administer the insulin? (Choose all that apply)	checkbox, Required
			0 | insulin_how___0 | Syringes 1 | insulin_how___1 | Pens with pen needles insulin_how___2 | Insulin pump 3 | insulin_how___3 | Other {insulin_how_oth} 4 | insulin_how___4 | Prefer not to answer
42	[initial_pump_cost] Show the field ONLY if: [insulin_how(2)] = '1'	How much did you pay out of pocket in [currency] for your [pump_type] pump? *Number only*	text (number, Min: 0) Custom alignment: LV
43	[pump_costs] Show the field ONLY if: [insulin_how(2)] = '1'	How much do you pay out of pocket in [currency] for a 1 month supply of insulin pump supplies? *Number only*	text (number, Min: 0) Custom alignment: LV

```
    id insulin_intake_pump usd_pump_unit usd_pumpsupplies_month
 <dbl>     <lgl>                   <dbl>                  <dbl>
   807     FALSE                     100                     45
```

▸ Same problem:
- one respondent could indicate pump-related OoPEs (#42 & #43) without having declared using a pump as a insulin administration method (#22)
- variable names are different between Q42-43 (*"insulin_how(2)"*) and Q22 (*"insulin_how__2"*).

▸ Note that it's also ID 807 as in the previous slide.