

1 **Analyzing open-ended questions in research: A commonly used category**
2 **selection methodology**

3 Luis D. Agosto Arroyo¹, Angerica Fitzmaurice⁶, Zlatan Feric⁶, David Kaeli⁶, John
4 Meeker⁵, Carmen Velez Vega¹, Akram Alshawabkeh³, José F. Cordero⁴, Nancy R.
5 Cardona-Cordero²

6 **Affiliations:**

- 7 1. School of Public Health, Medical Sciences Campus, University of Puerto Rico,
8 San Juan, PR
9 2. Department of Gynecology and Obstetrics, University of Rochester, Rochester,
10 NY
11 3. Department of Engineering, Northeastern University, Boston, MA
12 4. Department of Epidemiology and Biostatistics, College of Public Health,
13 University of Georgia, Athens, GA
14 5. School of Public Health, University of Michigan, Ann Arbor, MI
15 6. Department of Computer Science, Northeastern University, Boston, MA

16 *Nancy R. Cardona-Cordero, corresponding author

17 Nancy.cardona@upr.edu

18 **Contributions:**

19 Luis D. Agosto Arroyo - Conceptualization, Data curation, Formal analysis, Investigation,
20 Methodology, Supervision, Writing – original draft, Writing – review & editing
21 Angerica Fitzmaurice - Data curation, Investigation, Writing – review & editing
22 Zlatan Feric - Data curation, Investigation, Methodology
23 David Kaeli - Funding acquisition, Investigation, Supervision

24 John Meeker - Funding acquisition, Investigation, Project administration
25 Carmen Velez Vega - Funding acquisition, Investigation, Writing – review & editing
26 Akram Alshawabkeh - Conceptualization, Funding acquisition, Investigation, Project
27 administration, Supervision, Writing – review & editing
28 José F. Cordero - Conceptualization, Funding acquisition, Investigation, Project
29 administration, Supervision, Writing – review & editing
30 Nancy R. Cardona-Cordero - Conceptualization, Funding acquisition, Investigation,
31 Methodology, Project administration, Resources, Supervision, Writing – review & editing

32 **Abstract**

33 A closer examination of consumer product brands and how they are associated
34 with levels of potential endocrine disrupting chemicals should be explored. The large
35 number of brands available and changes in consumer preferences for certain brands
36 makes it difficult to develop questionnaires that include all brands. Open-ended brand
37 reporting questions are an option, but they bring challenges in identifying each brand
38 given the multiple possibilities of variations in brand name reporting. We report a
39 method for transforming product brand data reported as text to brand codes that allows
40 quantitative analysis of brand use and its association with endocrine disrupting
41 chemicals. We selected 14 consumer products to be included in our analyses. To
42 evaluate commonly used brand selection, we used Cohen's power calculations for two-
43 sample t-tests in R (version 1.3.0). Considering a moderate effect size (Cohen's d) of
44 0.5, each test will include the most used brand and the least used brand among the
45 commonly used brands per product and visit. We compared how the commonly used
46 brand selection differ per product in terms of the number of brands it selected, the total

47 sample size and the power calculated by creating a correlation matrix and analyzing the
48 relationship between power, commonly used brands, and brand usage. The correlation
49 coefficient between the commonly used brand frequency of each visit approximated
50 0.99. From all products, fabric softener, conditioner, and lotion were the products that
51 attained the highest power. The differences in brand use distributions per product
52 provided an optimal environment for evaluating the performance of the commonly used
53 brand selection methodology. It provides enough flexibility when selecting exposure
54 groups that it could be applied to any open-ended questions, and it proves significantly
55 useful when accounting for repeated measures.

56

57 **Introduction**

58 Numerous brands are being developed continuously for the diverse personal
59 consumer products marketed globally. These are products often used daily, weekly, or
60 monthly, and use of different brands may vary subject to the product availability, a
61 person's socio demographic characteristics. Use of personal care or home cleaning
62 products may contain fragrances and other additives that may expose people to harmful
63 chemicals such as phthalates and other persistent organic pollutants (POPs). Studies
64 have shown that consumer products can act as predictors of exposure for phenols,
65 parabens, phthalates, and many other chemicals (1-3). These substances are also
66 endocrine disrupting chemicals (EDCs) and have been associated with adverse
67 reproductive health outcomes (4-6). A recent study (7) found that elevated urine levels
68 of triclocarban and triclosan metabolites among participants who reported use of
69 specific bar soap brands. While several brands listed triclocarban/triclosan as an

70 additive in their bar soap, the levels were not listed and pointed out the importance of
71 examining specific brands rather than considering all brands of bar soap together. This
72 finding has led to a closer examination of consumer product brands and how they are
73 associated with levels of potential EDCs in the environment. The multiple brands of
74 commercial products present a major challenge to surveys because brand names vary
75 locally, regionally, and globally. Developing surveys with a systematic list of personal
76 care brands is complex and often not practical. We have used open ended questions in
77 our product use survey and extracted the brand name. We are unaware that this
78 approach has been used before in studies to assess human exposures and their
79 potential health effects.

80

81 A challenge to this type of analysis is that the large number of brands available
82 and changes in consumer preferences for certain brands makes it difficult to develop
83 questionnaires that include all brands. Moreover, long lists of brands complicate the
84 administration of questionnaires. Open-ended brand reporting is an option, but it brings
85 challenges in how to identify each brand given the multiple possibilities of variations in
86 brand name reporting. Moreover, for some personal care products there are many
87 brands are available with varying degrees of market share which requires an
88 examination of brands mostly used by consumers responding to the survey. Open-
89 ended responses pose a major challenge when studies are aiming to target specific
90 exposures in health sciences. For decades, closed questions have dominated
91 psychological, social and health sciences for obvious reasons including a generally
92 high-test score reliability as well as the possibility to standardize data collection and

93 study comparisons (8). However, open ended prompts may be reliably coded and
94 provide a more in-depth understanding of the exposure phenomena. Open questions
95 may instead be used as predictors in statistical analyses rather than endpoint variables,
96 providing and individual or collective variability in an exposure assessment of a possible
97 salient issue (8, 9).

98

99 We report a method for transforming product brand data reported as text into
100 brand codes that allows quantitative analysis of brand use and its association with
101 EDCs levels and its possible association with a health condition. We also describe an
102 approach to develop categories of most used brands, that account to majority of
103 exposures, when there are many brands involved.

104

105 **Methods**

106 **Data source**

107 We used data obtained by the Puerto Rico Testsite for Exploring Contaminations
108 Threats (PROTECT) study. PROTECT is a Superfund Research Center that examines
109 the effects of POPs on adverse pregnancy outcomes through a prospective maternal-
110 child cohort that evaluates fate and transport of environmental contamination and its
111 relationship with adverse pregnancy outcomes, such as preterm birth, in the northern
112 karst area in Puerto Rico. Details of this cohort has been described elsewhere (10, 11).

113

114 Our datasets originate from the PROTECT's Product Use Survey, which study
115 participants fill out at three separate follow up pregnancy visits, occurring at 16-, 20- and

116 24-weeks' gestation (+/- two weeks). The survey consists of a series of questions
117 regarding the use of 14 different personal care and household products designed to
118 quantify possible exposure to chemical substances. In addition to the yes or no
119 questions that assess if a participant used a certain consumer product. Those
120 responding yes to a specific personal care product are asked to write all the specific
121 brand names used as an open-ended question and stored as a text variable.

122

123 **Data transformation**

124 We develop a process to transform open-ended questions into coded categorical
125 variables. The first step focused on fixing structural errors of the brand names, which
126 included naming conventions, misspellings, incorrect capitalization, and Spanish
127 nomenclature of brand names that often lead to mislabeled categories (e.g., "Brand"
128 could also be spelled "brand" or "BRAND"). To address these inconsistencies, we used
129 Pandas `unique()` function to visualize the unique and different naming conventions of
130 the brand names. Since these values have a string data type, we used string methods
131 to match and rename them to a single consistent brand name that resulted in a single
132 categorical variable that contained a unique code for all reported brand names for each
133 specific consumer product studied.

134

135 The second step addressed was the structural data characteristic that some
136 participants had multiple entries of brand names for each product they used, per visit.
137 Therefore, we needed to separate these multiple entries using the string `split()`
138 function that splits around a passed delimiter, of which our brand names included

139 comma, “and”, “y” (Spanish for “and”), and ampersand (&). The returned data frame
140 with all split strings was used to create an array of columns (e.g. SOAP97_1,
141 SOAP97_2, SOAP97_3, etc.) and the original variable (SOAP97) is dropped using the
142 `drop()` function. Afterwards, we created a variable to count how many brands each
143 participant report using at each of three scheduled visits, and we transformed this new
144 data frame into a one-hot encoding data frame using Pandas `get_dummies()` function
145 that converts the categorical variables into indicator variables (dichotomous variables
146 that represent if the participant used that particular brand). The result is a cleaner data
147 frame that is primed for analysis.

148

149 Once we cleaned and structured the data set, we conducted a preliminary
150 analysis of the data to understand the distribution of most commonly used brands
151 (Figure 2). We encountered a high number of brand names and a highly skewed brand
152 use distribution where a few brands had high use and accounted for most of the brand
153 use. Therefore, we developed a selection process to identify what we called commonly
154 used brands, defined as brands that retained their use rank throughout the three time
155 periods of data collection. We used four criteria to select commonly used brands: 1)
156 high total sample size (total number of participants that used all the brands selected), 2)
157 high sample size per brand (number of participants that used each brand), 3) high
158 number of brands per product and 4) brand usage remaining consistent throughout the
159 three visits. This commonly used brand selection helped mitigate the effect that
160 selecting all brands for the analysis would have on our sample size. By selecting a

161 brand with low sample size, any future test would be statistically inconclusive (Kang et.
162 al 2008).

163

164 **Method evaluation**

165 To evaluate commonly used brand selection, we used Cohen's power
166 calculations for two-sample t-tests as described in the `pwr.t2n.test` function from the
167 `pwr` package for R (version 1.3.0). Considering a moderate effect size (Cohen's d) of
168 0.5, each test will include the highest ranked brand and the lowest ranked brand among
169 the commonly used brands per product and visit. The highest ranked brand has the
170 highest use frequency in contrast with the lowest ranked brands, which has the lowest
171 use frequency from the selected brands. To better visualize the relationship of the
172 highest ranked brand and the lowest ranked brand, a ratio was calculated to measure
173 how much larger is the sample size from the highest ranked brand. We also compared
174 how the commonly used brand selection differ per product in terms of the number of
175 brands it selected, the total sample size and the power estimate by calculating a
176 correlation matrix and analyzing the relationship between power, commonly used
177 brands and brand usage.

178

179 **Results**

180 The Product Use Survey collected data for 19 different personal care products, of which
181 only 14 products had enough data to conduct the commonly used brand selection
182 (Table 1). The products without sufficient data were excluded from this analysis. The
183 product with the most brands was perfume with 331 different brands and the product

184 with the least number of brands was shaving cream with 44 brands. The overall average
 185 for all studied products was 115 brands per product. The range of commonly used
 186 brands selected was between 2 (fabric softener and hair spray) and 8 (deodorant).

187
 188 As seen in Table 1, the most used product was deodorant, averaging 1463 participants
 189 per visit, while the least used product was shaving cream with an average of 97
 190 participants that used it per visit. Coincidentally the product with the highest commonly
 191 used brands, deodorant, reported the highest percentage of selected brand use per visit
 192 at around 92%. In contrast, one of the products with the lowest commonly used brands,
 193 fabric softener, reported a high percentage of selected brand use per visit at around
 194 77.3% when compared to other products with more commonly used brands such as
 195 makeup; fabric softener reported around 69.4% of selected brand use with only six
 196 selected brands.

197

198 **Table1. Commonly used brands frequency by visit.**

Type	Product	Total			Visit 1		Visit 2		Visit 3	
		Commonly used brands	Not commonly used brands	Used brands	Commonly used	Not commonly used	Commonly used	Not commonly used	Commonly used	Not commonly used
Personal care	Bar soap	6	93	99	1216 (89.9%)	136 (10.1%)	1182 (89.9%)	133 (10.1%)	1032 (90.1%)	113 (9.9%)
	Conditioner	4	149	153	603 (69.3%)	267 (30.7%)	549 (66.6%)	275 (33.4%)	521 (71.3%)	210 (28.7%)
	Deodorant	8	71	79	1363 (91.9%)	120 (8.1%)	1497 (91.7%)	136 (8.3%)	1178 (92.5%)	96 (7.5%)
	Hair spray	2	74	76	243 (58.4%)	173 (41.6%)	234 (59.8%)	157 (40.2%)	220 (63%)	129 (37%)
	Liquid soap	6	98	104	1041 (88.4%)	137 (11.6%)	1101 (88.1%)	149 (11.9%)	928 (89.4%)	110 (10.6%)
	Lotion	3	209	212	441 (53.3%)	387 (46.7%)	465 (54.3%)	392 (45.7%)	397 (53.8%)	341 (46.2%)
	Makeup	6	102	108	343 (71.0%)	140 (29%)	306 (68.2%)	143 (31.8%)	290 (69%)	130 (31%)

	Nail polish	3	46	49	195 (81.6%)	44 (18.4%)	197 (82.8%)	41 (17.2%)	158 (83.2%)	32 (16.8%)
	Perfume	3	331	334	360 (35.7%)	647 (64.3%)	391 (38.7%)	620 (61.3%)	314 (35.8%)	563 (64.2%)
	Shampoo	6	146	152	689 (78.3%)	191 (21.7%)	638 (76%)	201 (24%)	586 (80.8%)	139 (19.2%)
	Shaving cream	3	41	44	60 (63.2%)	35 (36.8%)	62 (57.9%)	45 (42.1%)	54 (60%)	36 (40%)
Home cleaning	Cleaning detergent	6	68	74	677 (83.2%)	137 (16.8%)	635 (81.5%)	144 (18.5%)	620 (84.8%)	111 (15.2%)
	Fabric softener	2	44	46	571 (77.2%)	169 (22.8%)	534 (78%)	151 (22%)	485 (76.7%)	147 (23.3%)
	General use cleaner	6	74	80	406 (91.2%)	39 (8.8%)	439 (90.1%)	48 (9.9%)	398 (89.6%)	46 (10.4%)

199
200 Table 2 shows the ratio and power when evaluating commonly and not commonly used
201 brands by visit. From all products, fabric softener, conditioner, and lotion were the
202 products that attained the highest power with the lowest ratio. Makeup was the following
203 product with a high power in the first visit, but a score of 0.49 and 0.54 was attained for
204 the second and third visit.

205
206 **Table 2. Power analysis of two sample t-test comparing the highest ranked**
207 **commonly used brand and the lowest ranked commonly used brand by visit.**

Type	Product	Visit 1				Visit 2				Visit 3			
		Max	Min	Ratio	Power	Max	Min	Ratio	Power	Max	Min	Ratio	Power
Personal care	Bar soap	523	32	16.3	0.78	505	45	11.2	0.89	420	38	11.1	0.83
	Conditioner	268	73	3.7	0.97	254	54	4.7	0.91	223	62	3.6	0.93
	Deodorant	412	46	9	0.89	417	40	10.4	0.85	348	28	12.4	0.72
	Hair spray	202	41	4.9	0.83	193	41	4.7	0.83	172	48	3.6	0.86
	Liquid soap	479	21	22.8	0.61	498	20	24.9	0.59	424	12	35.3	0.4
	Lotion	220	60	3.7	0.93	260	47	5.5	0.88	221	50	4.4	0.89
	Makeup	79	26	3	0.59	79	19	4.2	0.49	72	23	3.1	0.54
	Nail polish	153	14	10.9	0.43	168	12	14	0.38	134	10	13.4	0.33
	Perfume	226	53	4.3	0.90	254	56	4.5	0.92	197	50	3.9	0.88
	Shampoo	264	41	6.4	0.84	250	34	7.4	0.78	221	37	6	0.8
Shaving cream	33	10	3.3	0.27	34	13	2.6	0.32	25	12	2.1	0.28	

Home cleaning	Cleaning detergent	342	29	11.8	0.73	294	36	8.2	0.81	293	37	7.9	0.82
	Fabric softener	438	133	3.3	0.99	408	126	3.2	0.99	354	131	2.7	0.99
	General use cleaner	171	4	42.8	0.17	156	4	39	0.17	152	6	25.3	0.22

208 Max, Sample size of highest ranked commonly used brand; Min, Sample size of lowest
 209 ranked commonly used brand; Ratio, Ration between Max and Min.

210

211 In 1S Table, the correlation coefficient between the commonly used brand frequency of
 212 each visit approximated 0.99. Similarly, there was a moderate positive relationship
 213 between unique commonly used brands and ratio with a coefficient of 0.44 for visit 1,
 214 0.42 for visit 2 and 0.47 for visit 3. However, we observed a moderate inverse
 215 relationship between the total unique brands and the total brand usage. Finally, we
 216 found a moderate inverse relationship ($r < -0.56$) between the ratio and the power for all
 217 three visits.

218

219 Discussion

220 The differences in brand use distributions per product provided an optimal environment
 221 for evaluating the performance of the commonly used brand selection methodology. The
 222 variance of power calculated is not attributed to number of commonly used brands
 223 selected, but to frequency of use of the selected brands themselves. We observed that
 224 products with similar total unique brand and total unique commonly used brands,
 225 produced very different power. Shaving cream and fabric softener had similar commonly
 226 used brands, total used brands and ratio, but vastly different commonly used brand
 227 frequency, thus leading to the gap in calculated power as presented in Table 2. Bar

228 soap and liquid soap also had similar commonly used brands but different commonly
229 used brand frequency, yet their calculated power only deviated 0.03.

230

231 The commonly used selection methodology selected a relatively low amount of
232 commonly used brand from some products like fabric softener ($n = 2$) and hair spray (n
233 $= 2$). This can be attributed to the difference in brand use distribution per product.
234 However, this difference in commonly used brands between products doesn't translates
235 to low sample size. The percentage of participants that used the commonly used brands
236 for fabric softener was 77.2% in the first visit. Similar percentage was observed in
237 makeup with 71% for the first visit and for that product the number of commonly used
238 brand was six. The power for makeup and fabric softener was 0.59 and 0.99
239 respectively as described in Table 2. This suggests that even if the methodology selects
240 only a small number of brands, it makes up for it in performance of future statistical
241 tests.

242

243 The lowest performer regarding power was general use cleaner with a power of less
244 than 22 through the three visits (Figure 1S). This product has similar characteristics of
245 use as cleaning detergent, expect for the ratio between the highest ranked brand and
246 the lowest ranked brand amongst the commonly used brands. During post-collection
247 analysis, selecting low sample size exposure groups provides difficulty in establishing
248 significant research outcomes, therefore, using the commonly used brand selection
249 methodology as is for this product might not yield noteworthy results. However, products
250 like perfume who performed well with a power of around 0.90 and six commonly used
251 brands selected, in terms of the commonly used brand frequency for the three visits it

252 was 36.7%; suggesting that the number of commonly used brand selected could be
253 increased without affecting the legitimacy of the outcomes.

254

255 This methodology can be easily applied to any data set that records brand use in
256 repeated measures and contribute to the development of analysis plans that involve
257 open-ended questions in product use questionnaires and any other exposure related
258 forms. However, this methodology cannot be applied to cross-sectional studies (one
259 data point per participant) since the algorithm requires multiple measures to establish
260 the framework of what is a commonly used brand. This limitation is compensated by the
261 fact that not having multiple measures, sampling issues are not present, therefore
262 allowing to select brands based on optimal sample size.

263

264 A limitation of this process is that with open-ended questions some responses include
265 the company's name rather than the brand used. This may cause inaccuracies
266 regarding the measurement of EDCs exposure for that specific brand since companies
267 often sell a type of personal care product under multiple brands. This phenomenon, as
268 explained by Berger (15), could lead to a skewed reporting towards companies with
269 high number of brands.

270

271 The commonly used brand selection methodology allows researchers to objectively
272 analyze their data and not select brands for analysis based on subjective or arbitrary
273 conditions. While ranked selection methods exist in other fields, the Borda count (16) in
274 political science for example, we found no literature on brand selection methodologies

275 for environmental exposure analysis. Further studies using different distributions are
276 needed to test the reproducibility of this methodology within public health and
277 environmental research.

278 **Competing Interests**

279 The authors have declared that no competing interests exist.

280 **Acknowledgements**

281 We are extremely grateful to all PROTECT study participants and their families. The
282 authors also thank the nurses and research staff who participated in cohort recruitment
283 and follow up, as well as the Federally Qualified Health Centers (FQHC) in Puerto Rico
284 that facilitated participant recruitment, including Morovis Community Health Center,
285 Prymed in Ciales, Camuy Health Services, Inc. and the Delta OBGyn Group in Manatí,
286 as well as the Manatí Medical Center and the Metro Pavía Hospital in Arecibo.

287

288 **Funding**

289 This research is supported by grants from the National Institute of Environmental Health
290 Sciences (P42ES017198) and the Environmental Influences on Child Health Outcomes
291 (5UH3OD023251 and 3UH3OD023251-05S1) ECHO is a nationwide research program
292 supported by the NIH, Office of the Director to enhance child health.

293

294 The research protocol was approved by the Ethics and Research Committees of the
295 University of Puerto Rico, the University of Michigan School of Public Health, University

296 of Georgia, and Northeastern University. All study activities were fully explained to each
297 participant before obtaining written consent for participation in the study.

298

299 **References**

300 1. Meeker J, Cantonwine D, Rivera-González L, Ferguson K, Mukherjee B, Calafat A et al.
301 Distribution, Variability, and Predictors of Urinary Concentrations of Phenols and Parabens
302 among Pregnant Women in Puerto Rico. *Environmental Science & Technology*.
303 2013;47(7):3439-3447.

304 2. Cantonwine D, Cordero J, Rivera-González L, Anzalota Del Toro L, Ferguson K, Mukherjee
305 B et al. Urinary phthalate metabolite concentrations among pregnant women in Northern Puerto
306 Rico: Distribution, temporal variability, and predictors. *Environment International*. 2014;62:1-
307 11.

308 3. Rodríguez-Carmona Y, Ashrap P, Calafat A, Ye X, Rosario Z, Bedrosian L et al.
309 Determinants and characterization of exposure to phthalates, DEHTP and DINCH among
310 pregnant women in the PROTECT birth cohort in Puerto Rico. *Journal of Exposure Science*
311 *& Environmental Epidemiology*. 2019;30(1):56-69.

312 4. Ferguson K, Rosen E, Rosario Z, Feric Z, Calafat A, McElrath T et al. Environmental
313 phthalate exposure and preterm birth in the PROTECT birth cohort. *Environment International*.
314 2019;132:105099.

315 5. Watkins D, Ferguson K, Anzalota Del Toro L, Alshawabkeh A, Cordero J, Meeker J.
316 Associations between urinary phenol and paraben concentrations and markers of oxidative stress
317 and inflammation among pregnant women in Puerto Rico. *International Journal of Hygiene and*
318 *Environmental Health*. 2015;218(2):212-219.

319 6. Eick S, Meeker J, Swartzendruber A, Rios-McConnell R, Brown P, Vélez-Vega C et al.
320 Relationships between psychosocial factors during pregnancy and preterm birth in Puerto Rico.
321 *PLOS ONE*. 2020;15(1):e0227976.

322 7. Ashrap P, Watkins D, Calafat A, Ye X, Rosario Z, Brown P et al. Elevated concentrations of
323 urinary triclocarban, phenol and paraben among pregnant women in Northern Puerto Rico:
324 Predictors and trends. *Environment International*. 2018;121:990-1002.

325 8. Friberg O, Rosenvinge J. A comparison of open-ended and closed questions in the prediction
326 of mental health. *Quality & Quantity*. 2011;47(3):1397-1411.

327 9. Geer J. Do Open-Ended Questions Measure "Salient" Issues?. *Public Opinion Quarterly*.
328 1991;55(3):360.

- 329 10. Rivera-Núñez Z, Ashrap P, Barrett E, Llanos A, Watkins D, Cathey A et al. Personal care
330 products: Demographic characteristics and maternal hormones in pregnant women from Puerto
331 Rico. *Environmental Research*. 2022;206:112376.
- 332 11. Ferguson K, Rosario Z, McElrath T, Vélez Vega C, Cordero J, Alshawabkeh A et al.
333 Demographic risk factors for adverse birth outcomes in Puerto Rico in the PROTECT cohort.
334 *PLOS ONE*. 2019;14(6):e0217770.
- 335 12. Kang M, Ragan B, Park J. Issues in Outcomes Research: An Overview of Randomization
336 Techniques for Clinical Trials. *Journal of Athletic Training*. 2008;43(2):215-221.
- 337 13. Faber J, Fonseca L. How sample size influences research outcomes. *Dental Press Journal of*
338 *Orthodontics*. 2014;19(4):27-29.
- 339 14. Creusen M, Schoormans J. The Different Roles of Product Appearance in Consumer
340 Choice*. *Journal of Product Innovation Management*. 2005;22(1):63-81.
- 341 15. Berger J, Draganska M, Simonson I. The Influence of Product Variety on Brand Perception
342 and Choice. *Marketing Science*. 2007;26(4):460-472.
- 343 16. Emerson P. The original Borda count and partial voting. *Social Choice and Welfare*.
344 2011;40(2):353-358.

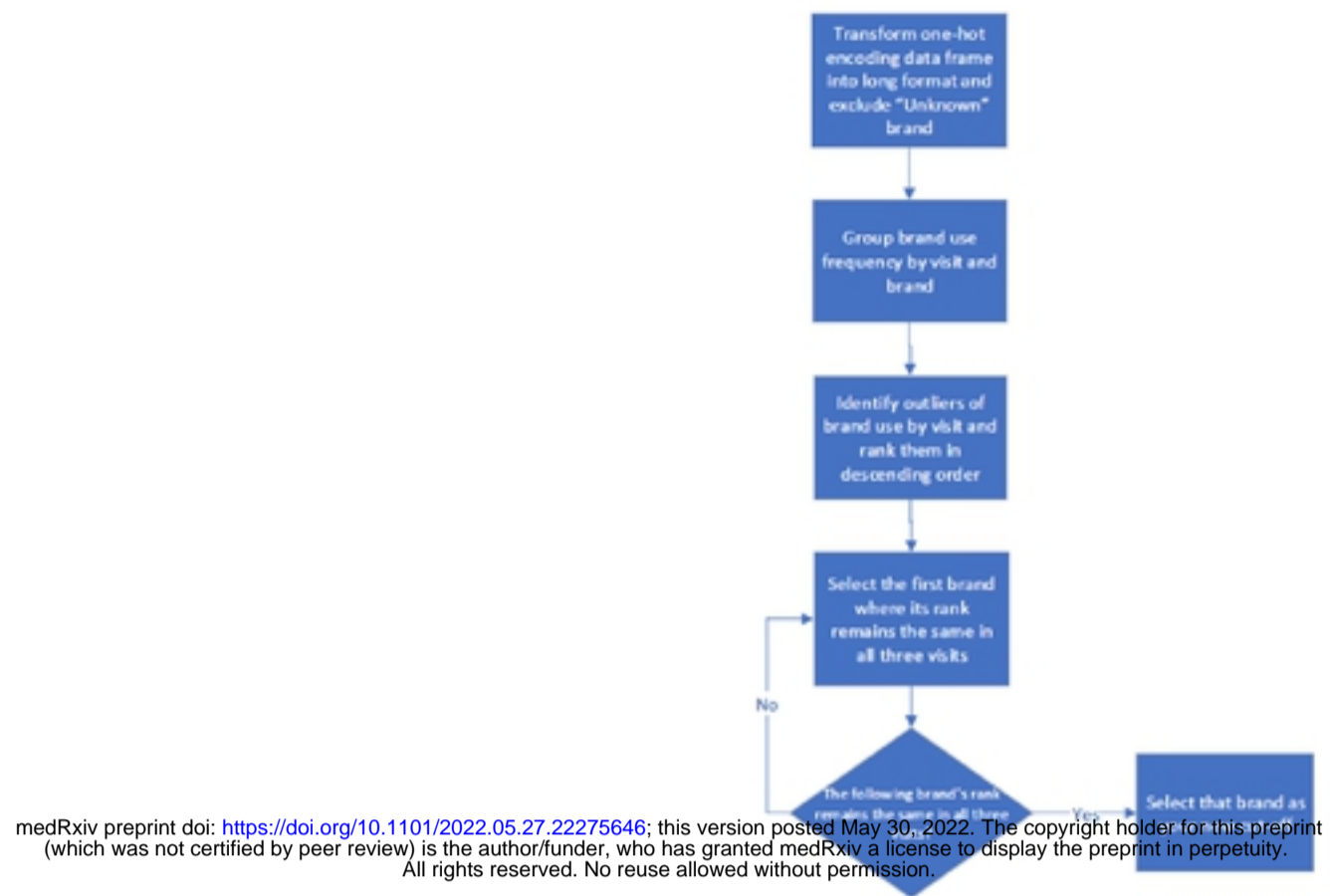


Figure 1. Commonly used brand selection flow chart.