
Patient-Focused Drug Development: Methods to Identify What Is Important to Patients

Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to <https://www.regulations.gov>. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Office of Communications, Division of Drug Information at druginfo@fda.hhs.gov, (855) 543-3784, or (301) 796-3400 or (CBER) Office of Communication, Outreach and Development at ocod@fda.hhs.gov, 800-835-4709 or 240-402-8010.

**U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)**

**October 2019
Procedural**

Patient-Focused Drug Development: Methods to Identify What Is Important to Patients Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders

*Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov*

<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>

and/or

*Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov*

<https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances>

**U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)**

**October 2019
Procedural**

Contains Nonbinding Recommendations

Draft — Not for Implementation

TABLE OF CONTENTS

| | | |
|-------------|--|-----------|
| I. | INTRODUCTION | 1 |
| A. | Overview of the Series of FDA Guidance Documents on Patient-Focused Drug Development | 1 |
| B. | Purpose and Scope of Guidance 2 | 2 |
| II. | METHODS TO IDENTIFY AND UNDERSTAND WHAT IS IMPORTANT TO PATIENTS | 3 |
| A. | Background Research | 3 |
| B. | Overview of Methods | 3 |
| III. | QUALITATIVE RESEARCH METHODS | 4 |
| A. | Common Qualitative Methods Used to Obtain Patient Input | 4 |
| 1. | <i>One-on-One Interviews</i> | 4 |
| 2. | <i>Focus Groups</i> | 6 |
| B. | Approaches to Asking the Right Questions | 8 |
| IV. | QUANTITATIVE RESEARCH METHODS | 9 |
| A. | Common Quantitative Methods Used to Obtain Patient Input | 9 |
| B. | Choice of Survey Administration Mode | 10 |
| C. | Considerations for Developing Items for a Survey Instrument | 10 |
| V. | MIXED METHODS | 15 |
| VI. | SPECIFIC POPULATIONS (CHILDREN, COGNITIVELY IMPAIRED, RARE DISEASES) AND CULTURALLY DIVERSE POPULATIONS | 18 |
| VII. | CONSIDERATIONS FOR USE OF SOCIAL MEDIA | 19 |
| | REFERENCES | 20 |
| | APPENDIX 1. Other Qualitative Methods | 23 |
| 1. | <i>Delphi methods</i> | 23 |
| 2. | <i>Observational methods</i> | 23 |
| 3. | <i>Facilitated discussions at patient meetings</i> | 23 |
| 4. | <i>Survey instruments with open-ended questions</i> | 24 |
| | APPENDIX 2. Considerations for Selection of Qualitative Data Collection Methods | 24 |
| | APPENDIX 3. Study Materials for Qualitative Studies | 27 |
| | APPENDIX 4. Analysis of Qualitative Data | 30 |
| | APPENDIX 5. Screening and Exit Interview Studies/ Survey Studies | 35 |
| | APPENDIX 6. Glossary | 37 |

Contains Nonbinding Recommendations

Draft — Not for Implementation

TABLE OF TABLES

| | |
|---|-----------|
| Table 1. Interview Types | 5 |
| Table 2. Some Potential Strengths and Limitations for Different Focus Group Modes | 7 |
| Table 3. Some Potential Strengths and Limitations of Open- and Closed-ended Questions | 13 |
| Table 4. Response Options | 13 |
| Table 5. Some Potential Strengths and Limitations for Different Qualitative Data Collection Methods..... | 24 |
| Table 6. Study Materials | 27 |
| Table 7. Steps Typically Used for Data Analysis in Qualitative Studies..... | 30 |
| Table 8. Modes for Displaying Qualitative Data..... | 32 |
| Table 9. Concept Saturation Table Example | 34 |

Contains Nonbinding Recommendations

Draft — Not for Implementation

1 **Methods to Identify What Is Important to Patients**
2 **Guidance for Industry, Food and Drug Administration Staff, and**
3 **Other Stakeholders¹**
4

5
6 This draft guidance, when finalized, will represent the current thinking of the Food and Drug
7 Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not
8 binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the
9 applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible
10 for this guidance as listed on the title page.
11

12
13
14 **I. INTRODUCTION**

15
16 **A. Overview of the Series of FDA Guidance Documents on Patient-Focused Drug**
17 **Development**
18

19 This guidance (Guidance 2) is the second in a series of four methodological *patient-focused*
20 *drug development*² (PFDD) guidance documents³ that FDA is developing to describe in a
21 stepwise manner how stakeholders (patients, researchers, medical product developers and others)
22 can collect and submit *patient experience data*⁴ and other relevant information from patients and
23 caregivers to be used for medical product⁵ development and regulatory decision-making. The
24 topics that each guidance document will address are described below.
25

- 26 • Methods to collect patient experience data that are accurate and representative of the
27 intended patient population (Guidance 1)⁶
28

¹ This guidance has been prepared by the Office of New Drugs (Center for Drug Evaluation and Research (CDER)), in cooperation with the Center for Biologics Evaluation and Research (CBER), at the Food and Drug Administration.

² Words or phrases found in the [Glossary](#) appear in bold italics at first mention within the body text in this document.

³ The four guidance documents that will be developed correspond to commitments under section I.J.1 associated with the sixth authorization of the Prescription Drug User Fee Act (PDUFA VI) under Title I of the FDA Reauthorization Act of 2017. The projected time frames for public workshops and guidance publication reflect FDA's published plan aligning the PDUFA VI commitments with some of the guidance requirements under section 3002 of the 21st Century Cures Act (available at <https://www.fda.gov/downloads/forindustry/userfees/prescriptiondruguserfee/ucm563618.pdf>).

⁴ 21st Century Cures Act: <https://www.congress.gov/bill/114th-congress/house-bill/34>

⁵ A drug, biological product, or medical device.

⁶ See the draft guidance for industry, FDA staff, and other stakeholders *Patient-Focused Drug Development: Collecting Comprehensive and Representative Input* (June 2018). When final, this guidance will represent FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA guidance web page at <https://www.fda.gov/RegulatoryInformation/Guidances/default.htm>.

Contains Nonbinding Recommendations

Draft — Not for Implementation

- 29 • Approaches to identify what is most important to patients with respect to their experience
30 as it relates to burden of disease and burden of treatment (Guidance 2)
31
- 32 • Approaches to identify and develop methods to measure impacts in clinical trials
33 (Guidance 3)
34
- 35 • Methods, standards, and technologies to collect and analyze clinical outcome assessment
36 data for regulatory decision-making (Guidance 4)
37

38 Please refer to the **Glossary, Guidance 1**, and other FDA guidances⁷ for additional information.
39 Many existing FDA regulations, guidances, and other standards and requirements pertaining to
40 the capture/collection, transmission, processing, storage, archiving, retention, and submission of
41 data from clinical studies conducted to support a regulatory medical product application (e.g., an
42 investigational new drug application (IND), new drug application (NDA), or biologics license
43 application (BLA) or medical product labeling language also apply to patient experience data
44 generated in studies.
45

46 FDA encourages stakeholders to have early interactions with FDA and obtain feedback from the
47 relevant FDA review division when considering collection of patient experience data related to
48 the burden of disease and burden of treatment. FDA recommends that stakeholders engage with
49 the appropriate subject matter experts (e.g., patients, qualitative researchers, survey
50 methodologists, statisticians, psychometricians, patient preference researchers) when designing
51 and implementing studies to evaluate the burden of disease, burden of treatment, and
52 perspectives on treatment benefits and harms.
53

B. Purpose and Scope of Guidance 2

54
55
56 This guidance describes methods to identify what matters most to patients regarding burden of
57 disease and burden of treatment to guide medical product development, including endpoint
58 development. This document discusses methods for collecting patient experience data, but it
59 should not be viewed as providing detailed instructions on how to use particular methods or as a
60 substitute for engaging subject matter experts when undertaking the work described.
61

62 The methods described in this document can be used to elicit what is important to patients, which
63 may in turn help inform the selection or development of ***clinical outcome assessments*** (COAs)
64 and the generation and use of ***patient preference*** information. However, this guidance does not
65 address methods for collecting and analyzing COA data, a topic to be covered in later guidance

⁷ See FDA guidance for industry *Patient Preference Information—Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling* (August 2016)

Contains Nonbinding Recommendations

Draft — Not for Implementation

66 in this series. It also does not address methods for collecting and analyzing patient preference
67 information, which is addressed in other FDA guidances.⁸

68
69 In general, FDA’s guidance documents do not establish legally enforceable responsibilities.
70 Instead, guidances describe the Agency’s current thinking on a topic and should be viewed only
71 as recommendations, unless specific regulatory or statutory requirements are cited. The use of
72 the word *should* in Agency guidances means that something is suggested or recommended, but
73 not required.

74
75

76 **II. METHODS TO IDENTIFY AND UNDERSTAND WHAT IS IMPORTANT TO** 77 **PATIENTS**

78

79 **A. Background Research**

80

81 Research to understand what matters most to patients living with a disease to guide medical
82 product development should begin with a good baseline characterization of the disease and
83 currently available treatments. Before conducting studies in patients, which may involve their
84 caregivers as well, literature reviews and consultation with relevant subject matter experts should
85 be used to develop research questions and select appropriate methods to identify what matters
86 most to patients.

87

88 **B. Overview of Methods**

89

90 When planning research, consider whether the study sample and inclusion criteria reflect the
91 *target population* characteristics, and whether the methods used to elicit information from
92 patients are appropriate for the research objective and target population.⁹

93

94 ***Qualitative research methods*** (e.g., through interviews or focus groups), ***quantitative research***
95 ***methods*** (e.g., through survey instruments), or ***mixed-methods research*** (e.g., through open-
96 ended and fixed-response items in a survey instrument) can be used to identify what is important
97 to patients. These methods can be used either independently or complementarily. When selecting
98 an appropriate research method, FDA recommends carefully considering the research objectives:

99

- 100 • Qualitative research methods are typically used to obtain a deeper understanding of the
101 patient experience by generating in-depth information about the experiences,
102 perspectives, and feelings of patients and others, in their own words.

103

⁸ Issues related to patient-reported outcome measures and patient preference information are addressed in the following guidances for industry: (1) *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims* (December 2009) and (2) *Patient Preference Information—Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling* (August 2016).

⁹ FDA draft guidance for industry *Patient-Focused Drug Development: Collecting Comprehensive and Representative Input* (June 2018).

Contains Nonbinding Recommendations

Draft — Not for Implementation

- 104
- 105
- 106
- 107
- 108
- 109
- 110
- 111
- Quantitative research methods are characterized by the collection of quantifiable data (e.g., numerical data) and the application of statistical methods to summarize the collected patient experience data, to describe, compare, or relate measures of patient experience.
 - Mixed-methods research involves using both qualitative and quantitative methods to understand the patient experience.

112 Although this document includes distinct sections for qualitative and quantitative research
113 methods, many data collection methods may be used in either approach. For example, patient
114 interviews are commonly used to generate qualitative data, but they also may be used to generate
115 quantitative data. Similarly, patient surveys are commonly used to generate quantitative data but
116 may also be used to generate qualitative data with open-ended questions.

117

118

119 **III. QUALITATIVE RESEARCH METHODS**

120

121 **A. Common Qualitative Methods Used to Obtain Patient Input¹⁰**

122

123 Methods of data collection most commonly include one-on-one interviews and focus groups.
124 Other qualitative methods that may be considered for use are summarized in [Appendix 1](#).¹¹
125 Before selecting qualitative data collection methods, consider potential strengths and limitations
126 of each of these methods, which are discussed in [Table 5 in Appendix 2](#). There is no single
127 preferred method for all uses and research questions.

128

129 *1. One-on-One Interviews*

130

131 One-on-one interviews involve a discussion on the topic of interest between the research
132 participant and a trained interviewer. Interviews offer opportunities to explore topics in depth at
133 an individual level using probing questions. The method is also used for exploring subject areas
134 that might be too sensitive for a *focus group* setting.

135

136 [Table 1](#) summarizes several interview types.

137

¹⁰ Research involving access to patient information or directly engaging with patients requires careful consideration of federal, state and local laws and institutional policies for the protection of human subjects. For additional details on human subjects protection, refer to [section IV.A.2 of Guidance 1](#).

¹¹ Many of these methods can be accomplished in person or using technology (e.g., social media, online forums, and web-based).

Contains Nonbinding Recommendations

Draft — Not for Implementation

138 **Table 1. Interview Types**

| Interview Method | Description |
|--|--|
| Semi-structured interviews ¹² | <ul style="list-style-type: none">• Interviewer leads the discussion using a semi-structured interview guide with standardized questions.• Interviewer can elicit further information (probe) about predetermined topics and their relative importance based on participant’s responses.• Leverages both open-ended questions as well as prompt and probing questions based on the specific research objective. |
| Structured interviews | <ul style="list-style-type: none">• Interviewer uses a structured interview guide with standardized questions.• The same questions should be asked of all participants, with no deviation or accommodation for participants’ responses.• Facilitates faster interviews that can be more easily analyzed and compared. |
| Unstructured interviews | <ul style="list-style-type: none">• Not led by predetermined questions.• The dialogue between the interviewer and participant remains open to the emergent priorities of the participant within the conversation. During the discussion, the interviewer provides little direction toward an a priori research agenda.• More time-consuming in the analysis phase than other methods and may not be ideal for capturing information targeted toward specific research questions. |

139

140

When the interview type and method of administration are determined, researchers should generally consider the following:

141

142

143

- Estimate the number of interviews to conduct.

144

145

- A greater number may be needed for unstructured or variable interviews, broad or complex topics, or heterogeneous populations

146

147

148

- Design interview questions and interview guide (focus on concepts of importance for *context of use* and research objectives).

149

150

151

- Pilot test interview guide (i.e., administer the interview in a small number of participants to identify and correct any methodological or logistical issues before using in the qualitative study).

152

153

154

155

- Select and train interviewers, considering expertise.

156

157

- Select sites to recruit participants (number of sites, geographic and patient representation).

158

159

160

FDA does not have a single recommended interview mode for eliciting patient input; however, an appropriate interview mode should be selected for the target population, study characteristics,

161

¹² Widely used method in qualitative research.

Contains Nonbinding Recommendations

Draft — Not for Implementation

162 and study objective(s). Each mode has strengths and potential limitations. For example, in-
163 person interviews can be conducted in a controlled environment, but logistical constraints (e.g.,
164 travel or time away from work or school) can limit participation. Telephone and video
165 conferences can provide an opportunity for including patients who would otherwise not be able
166 to participate in an in-person interview because of constraints related to travel and level of
167 patient impairment. A growing body of literature suggests no marked difference between the
168 modes of interviewing outlined above regarding the accuracy of data collected.¹³

169
170 Generally, qualitative researchers conduct and analyze interviews until **concept saturation** is
171 achieved (see [Appendix 4](#)).

172 173 2. *Focus Groups*

174
175 Focus groups involve a discussion with a group of participants (e.g., 5 to 10 participants) led by a
176 moderator. The moderator can explore issues both at the individual level and by encouraging
177 discussions among participants, which allows understanding of a range of experiences. Focus
178 groups can be conducted both in person or via telephone or online.

179
180 [Table 2](#) lists some potential strengths and limitations for each mode.
181

¹³ Block & Erskine, 2012; Cachia & Millward, 2011; Shapka et al., 2016; Vogl, 2013.

Contains Nonbinding Recommendations

Draft — Not for Implementation

182 **Table 2. Some Potential Strengths and Limitations for Different Focus Group Modes**

| Focus Group Mode | Strengths | Limitations |
|--|--|--|
| In-person focus groups | <ul style="list-style-type: none"> • Allows for collection of both verbal and nonverbal responses to help inform data interpretation. • A variety of written and brainstorming exercises (e.g., ranking exercises for concept of interest) can be incorporated into the study protocol to help elicit information. | <ul style="list-style-type: none"> • Cost can be prohibitive (e.g., moderator and participant travel costs; focus group facility rental fees). • Geographic restrictions can limit sampling pool to local participants. |
| Telephone or online focus groups (synchronous [takes place in real time] or asynchronous [takes place any time at convenience of the participant]) | <ul style="list-style-type: none"> • Participation is not limited to a geographic location – study sampling can be nationwide or worldwide. • Participants can be involved in the comfort of their homes or location of their choice. • Potential for participants to see each other if using a web cam – thus, allowing the potential benefits of seeing facial expressions. • No travel costs or focus group facility rental fees. | <ul style="list-style-type: none"> • Participants need to have access to the appropriate technology (e.g., a computer, telephone, webcam, internet service). • May be difficult to establish rapport between the interviewer and participant. • Participants may not have a private space to feel comfortable participating. • Disruptions (e.g., background noise and presence of family members) can interfere with sound quality and cause distractions. • Technical problems (e.g., wireless internet signal problems). |

183
184 Some important considerations for focus groups include:

- 185 • The number of focus groups to conduct, which may vary based on factors such as:
 - 186 ○ Complexity of the topic(s) (e.g., all versus some impacts of a disease on multiple
 - 187 ○ dimensions of a patient’s quality of life; dependent upon therapeutic area and
 - 188 ○ research question(s))
 - 189 ○ Heterogeneity of the participant sample
 - 190 ○ Number of *subgroups* planned (e.g., different age groups, disease severity
 - 191 ○ groups)
 - 192 ○ Initial data evaluation that suggests additional sessions are necessary to cover
 - 193 ○ topics sufficiently given the heterogeneity of the patients and themes and concepts
 - 194 ○ elicited
 - 195
 - 196
 - 197
 - 198
 - 199

Contains Nonbinding Recommendations

Draft — Not for Implementation

200 • Sample size for each focus group: generally, the goal is to keep the group small enough
201 to enable the elicitation of in-depth responses from each participant but large enough to
202 get a wide variety of perspectives across different severity levels and demographic
203 representation within the target disease.

204
205 ○ Although there is no set number recommended for a focus group sample, sample
206 sizes between 5 and 10 patients are common.

207
208 ○ A group may become fragmented (e.g., multiple, simultaneous conversations
209 occur) when there are too many participants, decreasing the likelihood of
210 engagement and responses from each individual.

211 212 **B. Approaches to Asking the Right Questions**

213
214 Regardless of method, the way questions are framed is critical to collecting unbiased patient
215 input. Although spontaneous responses are ideal, there are situations in which participants may
216 need to be prompted.

217
218 Prompts (i.e., open-ended questions to stimulate and provoke a participant’s memories) are used
219 to help the interviewer/moderator gain more information, particularly if the participant does not
220 initially provide detailed responses. However, consider the wording of prompt questions to avoid
221 leading the participant. Leading questions (i.e., questions that include or imply the desired
222 answer to the question in the phrasing of the question itself) are problematic because they may
223 result in biased or false/misleading answers (results). They may also lead to a missed opportunity
224 to hear an unexpected insight.

225
226 Approaches to *avoid* leading questions include (but are not limited to):

227
228 • Use a semi-structured interview guide with a set of prepared questions that act as a guide
229 to help facilitate discussion; additional probing based on the direction of the conversation
230 may be appropriate.

231
232 • Do not suggest an answer.

233
234 • Do not assume you know what the participant is thinking or feeling.

235
236 • Do not ask questions that cast judgment on a participant’s belief, choice, or perspective,
237 or imply that you prefer the participant to respond in one way versus another.

238
239 The boxed text that follows offers some examples of probing questions or prompts that are
240 leading or otherwise problematic.

241

Contains Nonbinding Recommendations

Draft — Not for Implementation

EXAMPLES

Example 1

Research Objective:

Determine what aspects of peripheral artery disease patients would like to see improved with treatment.

Leading Probing Question:

Wouldn't you consider it most important to improve your walking distance, for example, how far you walk around the track when you exercise?

Problem: This question guides the respondent to provide an answer that is more favorable or preferred by the researcher. Additionally, the *for example* clause may not include relevant examples for the research participant.

Potential Solution: Consider rephrasing as:

Think about the impact of peripheral artery disease on you. What would you most like to see improve with treatment?

Example 2

Research Objective:

Determine what factors caregivers consider when deciding whether to treat their child's autism with medication.

Probing Question That Casts Judgment:

"Could you tell me why you are not treating your child's autism with medication?"

Problem: This question implies that the interviewer is potentially casting judgment on the participant's beliefs or choices.

Potential Solution: Consider rephrasing as:

What did you consider when deciding whether to treat your child's autism with medication?

242

243

244 IV. QUANTITATIVE RESEARCH METHODS

245

246 A. Common Quantitative Methods Used to Obtain Patient Input

247

248 Survey research methods are commonly used to collect quantitative data from patients and
249 relevant stakeholders. Refer to **Guidance 1** for considerations for data management, data
250 analysis, and reporting of survey data. In designing a survey instrument, it is important to decide
251 how to administer the survey instrument and how to design and test the instructions, questions,
252 and response options.

253

Contains Nonbinding Recommendations

Draft — Not for Implementation

254 **B. Choice of Survey Administration Mode**

255
256 Survey instruments can be self-administered or interviewer-administered. Self-administered
257 surveys can be paper-based, telephone-based (e.g., interactive voice response system), or
258 electronic-based (e.g., computers, tablets, smartphone). Interviewer-administered surveys can be
259 conducted face-to-face or remotely (e.g., via telephone or web).

260
261 Choice of mode of administration may be driven by a variety of factors. For example, an
262 interviewer-administered survey instrument or a survey instrument using an interactive voice
263 response system may be useful in patients with visual impairment.

264
265 Self-administration allows participants to respond at their own pace and at their convenience.
266 Compared to interviewer-administered surveys, self-administration typically is less costly and
267 removes the potential for interviewer bias. Computer-administered survey instruments can assist
268 the respondent in navigating skip patterns in a survey, help minimize item-level missing data,
269 allow for faster data collection and analysis compared with other methods (e.g., paper-based
270 survey instruments), and allow real-time data analysis. Social media also might be used to
271 implement survey instruments. Best practices for designing and implementing studies using
272 survey instruments and technology also are applicable to the use of social media to conduct a
273 survey.

274
275 Using survey instruments in a clinical trial for screening and/or exit visits may add greater depth
276 to understanding the burden of disease and treatment, as well as provide more detail on patients'
277 perspectives on treatment benefits and harms (see [Appendix 5](#)).¹⁴

278
279 If the instrument is intended to be used to derive a study endpoint(s) in a clinical trial to support
280 labeling claims, FDA recommends that stakeholders refer to future PFDD guidance documents
281 regarding clinical outcome assessments to ensure the instrument is appropriate for use in this
282 context.

283 **C. Considerations for Developing Items for a Survey Instrument**

284
285
286 Survey instrument instructions and items should be:¹⁵

- 287
- 288 • Well-aligned with the research objective(s) and designed to answer the research questions
 - 289
 - 290 • Specific to the concept of interest (e.g., disease symptoms and impacts, current treatment,
 - 291 past treatments, treatment side effects)
 - 292
 - 293 • Well-understood by participants to enhance consistency of response, including:
 - 294

¹⁴ For more information on steps to follow when conducting survey instruments used in noninterventional studies, see Cooper et al, 2006.

¹⁵ Questions for survey instruments can be generated from multiple sources that can be found in standard text books (e.g., Streiner, Norman, & Cairney, 2015).

Contains Nonbinding Recommendations

Draft — Not for Implementation

- 295 ○ Assessed appropriateness for literacy and numeracy of the population
- 296
- 297 ○ Incorporated natural familiar language (e.g., minimal use of clinical terminology)
- 298 as used by patients when discussing the concept of interest
- 299
- 300 ○ Assessed translatability of items if using survey instruments in multinational and
- 301 multicultural studies; ultimately, the survey instrument should be translated and
- 302 culturally adapted for all languages and cultures where it will be administered
- 303
- 304 ○ Tested through interviews of respondents in the target population to make sure
- 305 they interpret the survey instrument instructions, items, and responses as intended
- 306 and can respond accordingly (including that item stems and response options are
- 307 appropriate and meaningful)
- 308
- 309 ● Formatted in a simple manner to maximize the ease of use for respondents and
- 310 interviewers
- 311
- 312 ● Tested for usability, if the survey instrument is electronic or web-based
- 313
- 314 ● Scripted to ensure standardization, if administered by an interviewer
- 315
- 316 ● Assessed for potential social desirability bias (i.e., the tendency of respondents to answer
- 317 questions in a manner they perceive may be viewed favorably by others)
- 318
- 319 ● Assessed for applicability of the content (although sometimes a *not applicable* response
- 320 option is also needed for an item)
- 321

322 The following question formats should generally be avoided:

- 323
- 324 ● Incomplete questions (e.g., Age? Reason last saw doctor?)
- 325
- 326 ● Poorly worded questions (e.g., poorly defined terms)
- 327
- 328 ● Double-barreled or multi-barreled questions (i.e., a question that asks about two or more
- 329 concepts at once)
- 330
- 331 ● Double negatives (i.e., a sentence that includes two negatives)
- 332
- 333 ● Leading questions
- 334

Contains Nonbinding Recommendations

Draft — Not for Implementation

EXAMPLE

Double-Barreled and Potentially Leading Question:

How embarrassed and self-conscious have you been because of your condition?

This question is asking about two different concepts:

1. How embarrassed have you been because of your condition?
2. How self-conscious have you been because of your condition?

Combining these concepts into one question makes it unclear about what is being measured. Once respondents answer the question, it likely will be impossible to know which concept the respondents were thinking about when they answered the question (unless it was an interviewer-administered question and further probing was done). The question may lead a respondent to report some degree of embarrassment or self-consciousness even though the respondent may feel neither embarrassed nor self-conscious.

Double-Negative Question:

Do you agree or disagree with the following statement? “Doctors should never be allowed not to discuss urgent lab results with patients on weekends.”

If you disagree, you are saying that you do not think that doctors ***should not*** discuss urgent lab results to patients on the weekends. In other words, you probably believe that doctors should discuss urgent lab results with patients on the weekends.

If a negative item is in fact needed for a survey instrument, you should underline the negative word or words to catch the participant’s attention. However, it is best to avoid the use of negative items, because some instruments may be administered via telephone interview or interactive voice response systems where there is no visually accessible (i.e., visible) question. Instruments should be developed so that they have the potential to be implemented across all modes of data collection.

335

336 There are two types of questions that can be used in survey instruments¹⁶:

337

- Closed-ended questions (questions with fixed set of response options)
- Open-ended questions (questions without a fixed set of responses options, e.g., free text)

338

339 **Table 3** lists examples of closed- and open-ended questions, as well as some strengths and
340 potential limitations of using different question types.

341

342

343

¹⁶ These types of questions also can be asked in interviews and focus group discussions.

Contains Nonbinding Recommendations

Draft — Not for Implementation

344 **Table 3. Some Potential Strengths and Limitations of Open- and Closed-ended Questions**

| Question Type | Examples | Strengths | Limitations |
|------------------------|--|--|--|
| Closed-ended questions | Which of the following health conditions do you currently have? <ul style="list-style-type: none"> • Asthma • Acne • High blood pressure • Glaucoma | <ul style="list-style-type: none"> • Respondent typically can more reliably answer the question when response options are given • Researcher typically can more reliably interpret answers • Easier and quicker for respondents to record answers | <ul style="list-style-type: none"> • May not provide respondent with a comprehensive list of response options • Response options may not be applicable to the respondent |
| Open-ended questions | What health conditions do you have? | <ul style="list-style-type: none"> • May obtain answers that were unplanned • May obtain more realistic answers • Provides opportunity for respondents to answer questions in their own words | <ul style="list-style-type: none"> • Less common answers may be challenging to analyze |

345
346 [Table 4](#) examines some different types of response options.

347
348 **Table 4. Response Options**

| Response Option | Examples | Considerations |
|------------------------------------|---|--|
| Checklist | Please check to indicate that you have ever had the following conditions (check all that apply): <ul style="list-style-type: none"> • Diabetes • Kidney disease • Stroke • High blood pressure • Asthma • Heart attack | <ul style="list-style-type: none"> • Checklists may not cover all the possible responses; in these instances, free text may be needed |
| Dichotomous (two response options) | Have you ever been diagnosed with glaucoma? <ul style="list-style-type: none"> • Yes • No I have been diagnosed with glaucoma <ul style="list-style-type: none"> • True • False | <ul style="list-style-type: none"> • May force respondents to choose between a narrow set of response options, resulting in a response that does not completely capture their experiences/feelings • Limits the analysis that can be performed |

(Table continued)

Contains Nonbinding Recommendations

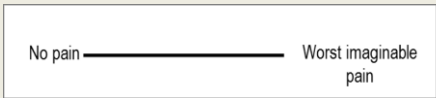
Draft — Not for Implementation

| Response Option | Examples | Considerations |
|------------------------|---|---|
| Rankings | <p>Please rank the importance of the following characteristics of a treatment for lung cancer (Fill in your rank order in the spaces provided using the numbers 1 through 5, with 1 indicating most important and 5 indicating least important).</p> <ul style="list-style-type: none"> — Treatment relieves symptoms — Treatment has few side effects — Treatment will increase survival — Treatment can be taken as a pill — Treatment can be taken monthly | <ul style="list-style-type: none"> • Ranking can be a difficult task for respondents, particularly if there are several response options (e.g., >5) and/or if respondents have poor numeracy skills • Rank order items can be difficult to relate to other variables • Ranking assumes that participants rate all options in a hierarchical fashion in order of importance when that may not be the case (e.g., a person might think only one of the options is truly important and the others are not important; a person may not differentiate among all the options) |
| Rating scales | <p><u>Numeric</u> Please rate your pain at its worst in the last 24 hours.</p> <ul style="list-style-type: none"> • 0 (no pain) • 1 • 2 • 3 • 4 • 5 • 6 • 7 • 8 • 9 • 10 (worst imaginable pain) <p><u>Verbal</u> Please rate your pain at its worst in the last 24 hours.</p> <ul style="list-style-type: none"> • None • Mild • Moderate • Severe <p>How often have you had pain during the past week?</p> <ul style="list-style-type: none"> • Not at all • A little • Quite a bit • All the time | <ul style="list-style-type: none"> • Decreased validity with extremes of age, e.g., young children (numeric) • Limited number of response categories (verbal) • Although distances between verbal descriptors on verbal rating scales appear equidistant, the actual observed distances may vary (verbal) • Susceptible to language/cultural effects and/or literacy effects (verbal, numeric) |

(Table continued)

Contains Nonbinding Recommendations

Draft — Not for Implementation

| Response Option | Examples | Considerations |
|---------------------------|--|--|
| Visual analog scale (VAS) | <p>How severe has your abdominal pain been today? (Place a mark (I) on the line below)</p>  | <ul style="list-style-type: none">• False sense of precision• Cannot be administered verbally• Higher rates of missing data (Dworkin et al., 2005; Hawker et al., 2011)• Inconsistencies with the length of VAS line• Marks may not be clear |

349

350 The ordering of questions in a survey instrument also may be important. The way a person
351 responds to a question can be influenced by earlier questions (e.g., priming can occur when
352 information presented in an earlier question causes respondents to adjust their responses to
353 subsequent questions).

354

355 Priming can be problematic in survey research. Ways to avoid priming include:

356

- Order questions to avoid influencing the answers given (i.e., question order bias).
- Use appropriate spacing of questions (separate topics on different pages or electronic screens).

357

358
359
360
361 In some instances, a screening question may be needed in a survey instrument to ensure the
362 survey instrument is appropriate and relevant to the respondent.

363

364

EXAMPLE

Scenario: A survey instrument has been designed to assess the burden of using a colostomy bag (stoma bag).

A screening question would be useful to avoid including survey respondents in which this subject matter may not be of relevance.

Screening question: *Do you currently use a stoma bag? Yes/No*

365

366

V. MIXED METHODS

367 *Overview of Mixed Methods (Research That Use Both Qualitative and Quantitative*
368 *Methods)*

369

370 Mixed-methods research involves using both qualitative and quantitative methods.

371 For additional details on mixed methods, refer to **Section III.C of Guidance 1**. FDA encourages
372 researchers to consider the goals and objectives of using a mixed-methods approach and how the
373

Contains Nonbinding Recommendations

Draft — Not for Implementation

374 results from both qualitative and quantitative research components are intended to be used
375 together.¹⁷

376

377 Questions researchers should ask when determining the rationale for using a mixed design:

378

379 • What is the goal for mixing quantitative and qualitative approaches (i.e., what do you
380 want to achieve)?

381

382 • How will a mixed-method approach help answer the research question(s)?

383

384 Reasons to use a mixed design may include:

385

386 • Harmonizing and confirming results from different methods (triangulation)

387

388 • Supplementing and clarifying results from one method with results from another method
389 (complementarity)

390

391 • Using results from one method to inform the design of another method

392

393 • Discovering inconsistencies, contradictions, and new perspectives, and reframing of
394 questions or results from one method with questions or results from the other method
395 (initiation)

396

397 • Expanding the scope (range) of a research question by using different methods for
398 different components of the research question (expansion)

399

400 Questions researchers should ask to determine what specific mixed design to use:

401

402 • Will qualitative or quantitative methods be more predominant in the study, or will both
403 be given equal status in the study?

404

405 • Should qualitative and quantitative components be carried out concurrently or
406 sequentially?

¹⁷ Johnson & Christensen, 2014.

Contains Nonbinding Recommendations

Draft — Not for Implementation

EXAMPLES

Mixed-method study based on a qualitatively driven concurrent design

A study examines the patient experience with living with amyotrophic lateral sclerosis via in-depth interviews with patients and caregivers (qualitative component). Within the study, there is concurrent collection of symptom checklist data (quantitative component). The data from both study components are analyzed separately before being compared.

Mixed-method study based on a quantitatively driven sequential design

A study explores depression and anxiety in patients with acute coronary syndrome (ACS) by administering a questionnaire to patients with ACS (quantitative component). Analyzing the questionnaire data, researchers find an association between depression and anxiety for female patients with ACS. In the second phase of the study, the researchers conduct follow-up qualitative interviews with a sample of the patients, enriching for the most depressed male and female patients with ACS, to explore why the relationship was present only for female patients (qualitative component).

Mixed-method study with equal status sequential design

A phase 2 clinical trial evaluates the efficacy of a medical product for the treatment of nontuberculous mycobacteria. Within the trial, symptom questionnaires are administered at baseline and follow-up visits. The treatment and control groups are compared on the quantitative data (quantitative component).

At the end of the clinical trial, qualitative interviews are sequentially conducted to obtain an in-depth understanding of the meaningfulness of the patients' experiences (e.g., input on relevance of questionnaire, meaningful symptom improvement) within the clinical trial (qualitative component). Approximately one-half of the treatment and control group members are interviewed and their responses are compared. The data from both study components are analyzed separately, and mixing takes place in interpreting the final phase 2 study results.

407
408 In some cases, a mixed-methods approach will be employed by first conducting qualitative
409 research to generate concepts and better understand an experience and/or event as expressed in
410 the patient's own voice, and subsequently using this information to develop a quantitative survey
411 to better understand the prevalence of those concepts in a larger patient population.¹⁸
412
413

¹⁸ For additional details on mixed methods, refer to [section III.C of Guidance 1](#). For other considerations and methods for operationalizing mixed-method studies, you may refer to Johnson & Christensen, 2014; Johnson & Christensen, 2017; and Teddlie & Tashakkori, 2009.

Contains Nonbinding Recommendations

Draft — Not for Implementation

414 VI. SPECIFIC POPULATIONS (CHILDREN, COGNITIVELY IMPAIRED, RARE 415 DISEASES) AND CULTURALLY DIVERSE POPULATIONS

416
417 *What are some special considerations when obtaining input from specific populations*
418 *(children, cognitive impairment, rare diseases)?*

- 419
- 420 • The patient’s health status (e.g., some populations may experience fatigue because of
421 illness or travel may affect the quality of data in a lengthy assessment).
422
 - 423 • Certain populations may have a limited attention span (e.g., young children) or cognitive
424 slowness (e.g., elderly).
425
 - 426 • Using nontraditional approaches to stimulate participation may be appropriate (e.g.,
427 asking young children to participate in drawing activities can help elicit concepts, use of
428 props).¹⁹
429
 - 430 • Remote assessment can be useful in geographically diverse patient populations.
431
 - 432 • The emotional burden of the respondent (potential for heightened emotions, including
433 anxieties and discomfort among patients and caregivers), as well as the emotional burden
434 of the interviewer (potential for emotional distress associated with hearing about difficult
435 patient and caregiver experiences) may affect responses.
436
 - 437 • The patients’ stage in their disease course, because understanding and acceptance of
438 prognosis change over time.
439
 - 440 • In cases where the patient can provide reliable self-reporting, consider whether to
441 conduct qualitative patient interviews with the caregiver present or absent. Generally, the
442 caregiver should not be present during the patient interview (e.g., they may be asked to sit
443 outside the room). In cases where it is important for the caregiver to be present with the
444 patient (e.g., for patient comfort), the caregiver could sit behind the respondent to
445 minimize influencing the interview (either verbally or nonverbally). In either case, the
446 protocol should include a clear plan for how data from the patient and caregiver will be
447 collected and reported so that the source of the information is clear for analysis and
448 reporting.
449
 - 450 • For patients who are unable to self-report, eliciting what behaviors caregivers observe in
451 the patients (including things the patients tell them) can help to avoid proxy reporting
452 (i.e., reporting from the caregiver as if they were the patient). Proxy reporting can lead to
453 inappropriate inferences and may not be reflective of what a patient may be truly thinking
454 or feeling.
455

¹⁹ For additional information on concept elicitation in children see Matza, Patrick, Riley et al., 2013.

Contains Nonbinding Recommendations

Draft — Not for Implementation

456 ***What are some special considerations when obtaining input from patients from different***
457 ***cultures?***
458

- 459 • In both qualitative and quantitative studies, translation and cultural adaptation procedures
460 for multinational, multiregional, and/or multicultural survey studies should generally be
461 used to keep the meaning of questions similar. In survey instruments, it is generally
462 helpful to also keep the format of the questions similar across translations, considering
463 the limits of the target language, and to retain the properties of the instrument such as
464 range of response options and scoring.
465
- 466 • Poorly translated surveys can prevent researchers from collecting data comparable to that
467 of surveys in the source (original) language.²⁰ Ideally, translatability assessment should
468 be performed early during development of a survey instrument to address the needs of
469 different nationalities, regions, and cultures.²¹
470

471 **VII. CONSIDERATIONS FOR USE OF SOCIAL MEDIA**

472

473 Data collected prospectively using social media could be derived from qualitative, quantitative,
474 or mixed methods. Other research designs such as mixed-method sequential research designs can
475 further strengthen the depth of knowledge gained from social media research.
476

477 The researcher should carefully select the source(s) of the social media with the research
478 question in mind, because findings across social platforms may be distinctly different (e.g.,
479 certain platforms may have strong advocacy/support community presence; others may
480 predominantly capture industry/academic perspectives surrounding certain issues). Different
481 social media communities appeal to different segments of the population, and a community's
482 degree of user anonymity may affect what users are willing to discuss. When possible, social
483 media research should examine a variety of social media networks and communities to obtain
484 data that can be most generalized to the population of interest.²² A discussion of the strengths
485 and limitations of using social media in qualitative, quantitative, and mixed methods research
486 can be found in Guidance 1.
487

488 Ideally, research examines data from communities that provide personal information (e.g.,
489 verified patient communities) to allow verification of personal characteristics. However, in some
490 cases, it may be appropriate to examine communities that allow users to remain anonymous or
491 post under a username (e.g., blogs and forums), particularly when topics are of a sensitive nature.
492

493 When considering social media data analysis, data collection methods should address potential
494 limitations (e.g., lacking mechanisms to verify patient characteristics, such as identity, diagnosis,
495 or other patient characteristics) and how these limitations can affect data integrity and
496 interpretation.

²⁰ Survey Research Center, 2016; Wild et al., 2005.

²¹ Further information on translation and culturability adaptation of survey instruments can be found in Survey Research Center, 2016; Wild et al., 2005.

²² It is important to consider ethical standards (e.g., disclosure, consent, data ownership) for the collection and analysis of social media data. For a discussion on ethical considerations, refer to Gleibs, 2014.

Contains Nonbinding Recommendations

Draft — Not for Implementation

REFERENCES

- 497
498
499 Block ES, Erskine L, “Interviewing by Telephone: Specific Considerations, Opportunities, and
500 Challenges.” *International Journal of Qualitative Methods*, vol. 4, pp. 428-445, 2012.
- 501 Cachia M, Millward L, “The Telephone Medium and Semi-Structured Interviews: A
502 Complementary Fit.” *Qualitative Research in Organizations and Management: An
503 International Journal*, vol. 6(3), pp. 265-277, 2011.
- 504 Cooper CJ, Cooper SP, Del Junco DJ, et al., “Web-based Data Collection: Detailed Methods of a
505 Questionnaire and Data Gathering Tool.” *Epidemiologic Perspectives & Innovations*, vol.
506 3(1), p. 1, 2006.
- 507 Dworkin RH, Turk DC, Farrar JT, et al., “Core Outcome Measures for Chronic Pain Clinical
508 Trials: IMMPACT Recommendations.” *Pain*, vol. 113, pp. 9-19, 2005.
- 509 FDA, guidance for industry, *Patient Preference Information –Voluntary Submission, Review in
510 Premarket Approval Applications, Humanitarian Device Exemption Applications, and De
511 Novo Requests, and Inclusion in Decision Summaries and Device Labeling*, August 2016.
512
- 513 FDA, guidance for industry, *Patient-Reported Outcome Measures: Use in Medical Product
514 Development to Support Labeling Claims*, December 2009.
- 515 FDA-NIH Biomarker Working Group, “BEST (Biomarkers, EndpointS, and other Tools)
516 Resource,” Glossary, Published January 28, 2016 (updated May 2, 2018), available at
517 <https://www.ncbi.nlm.nih.gov/books/NBK338448>.
- 518 Gibbs GR, *Qualitative Research Kit: Analyzing Qualitative Data*, London, England: SAGE
519 Publications Ltd., 2007.
- 520 Gleibs IH, “Turning Virtual Public Spaces Into Laboratories: Thoughts on Conducting Online
521 Field Studies Using Social Network Sites. *Analyses of Social Issues and Public Policy*, vol.
522 14, pp. 352–370, 2014.
- 523 Grajales FJ 3rd, Sheps S, Ho K, et al., “Social Media: A Review and Tutorial of Applications in
524 Medicine and Health Care.” *Journal of Medical Internet Research*, vol. 16(2): p. e13, 2014.
- 525 Hawker GA, Mian S, Kendzerska T, et al., “Measures of Adult pain: Visual Analog Scale for
526 Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire
527 (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale
528 (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and
529 Constant Osteoarthritis Pain (ICOAP).” *Arthritis Care & Research*, vol. 63, Supplement 11,
530 pp. S240-252, 2011.
- 531 Institute of Medicine Committee on Understanding the Biology of Sex and Gender Differences,
532 *Exploring the Biological Contributions to Human Health: Does Sex Matter?* TM Wizemann
533 and ML Pardue, editors. Washington, DC: National Academies Press, 2001.
- 534

Contains Nonbinding Recommendations

Draft — Not for Implementation

- 535 International Conference on Harmonisation (ICH) Guideline, *Revision of M4E Guideline on*
536 *Enhancing the Format and Structure of Benefit-Risk Information in ICH* (Efficacy –
537 M4E(R2)), June 2016, available at
538 http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4E_R2_Efficacy/M4E
539 [_R2__Step_4.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4E_R2_Efficacy/M4E_R2__Step_4.pdf).
- 540 Johnson, RB, and Christensen, L, *Educational Research*, Thousand Oaks, CA: SAGE
541 Publications, Inc., 2014.
- 542 Johnson, RB, and Christensen, L, *Educational Research: Quantitative, Qualitative, and Mixed*
543 *Approaches*, Thousand Oaks, CA: SAGE Publications, Inc., 2017.
- 544 Keeney S, McKenna H, and Hasson F, *The Delphi Technique in Nursing and Health Research*,
545 John Wiley & Sons, 2010.
- 546 Matza LS, Patrick DL, Riley AW, et al., “Pediatric Patient-Reported Outcome Instruments for
547 Research to Support Medical Product Labeling: Report of the ISPOR PRO Good
548 Research Practices for the Assessment of Children and Adolescents Task Force.” *Value*
549 *in Health*, vol. 16(4), pp. 461-479, 2013.
- 550 Medical Device Innovation Consortium. “Medical Device Innovation Consortium (MDIC)
551 Patient Centered Benefit-Risk Project Report: A Framework for Incorporating
552 Information on Patient Preferences Regarding Benefit and Risk into Regulatory
553 Assessments of New Medical Technology,” April 22, 2015.
- 554 Patrick DL. “What Does “Patient-Centered Outcomes” Mean?” 18th Annual International
555 Meeting of the International Society for Pharmacoeconomics and Outcomes Research
556 Plenary Session, May 2013.
- 557 Public Law 102-73. The National Literacy Act of 1991.
- 558 Shapka JD, Domene JF, Khan S, et al., “Online Versus In-Person Interviews With Adolescents:
559 An Exploration of Data Equivalence.” *Computers in Human Behavior*, vol. 58, pp. 361-367,
560 2016.
- 561 Stanford University Libraries n.d.(b), “About Data Management Plans (DMPS),” available at
562 <https://library.stanford.edu/research/data-management-services/data-management-plans>.
- 563 Streiner DL, Norman, GR, Cairney J, *Health Measurement Scales: A Practical Guide to Their*
564 *Development and Use* (5th ed.), New York, NY: Oxford University Press, 2015.
- 565 Survey Research Center. (2016). “Guidelines for Best Practice in Cross-Cultural Surveys.” Ann
566 Arbor, MI: Survey Research Center, *Institute for Social Research*, University of Michigan at
567 Ann Arbor,. Retrieved July 03, 2018 (available at <http://www.ccsr.isr.umich.edu/>).
- 568 Teddlie C, and Tashakkori A, *Foundations of Mixed Methods Research: Integrating Quantitative*
569 *and Qualitative Approaches in the Social and Behavioral Sciences*, Thousand Oaks, CA:
570 SAGE Publications, Inc., 2009.
- 571 Tong A, Sainsbury P, Craig J, “Consolidated Criteria for Reporting Qualitative Research
572 (COREQ): A 32-item Checklist for Interviews and Focus Groups.” *International Journal for*
573 *Quality Health Care*, vol. 19(6), pp. 349-357, 2007.

Contains Nonbinding Recommendations

Draft — Not for Implementation

- 574 U.S. Department of Health and Human Services. “Healthy People 2010”. Washington, DC: U.S.
575 Government Printing Office. Originally developed for Ratzan SC, Parker RM. 2000.
576 Introduction. In National Library of Medicine Current Bibliographies in Medicine: Health
577 Literacy. Selden CR, Zorn M, Ratzan SC, Parker RM, Editors. NLM Pub. No. CBM 2000-1.
578 Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human
579 Services, 2000.
- 580 University of California San Francisco.. “Clinical Trial Protocol Development,” October 3, 2017.
581 Retrieved August 07, 2019, from <https://hub.ucsf.edu/protocol-development>
- 582 Vogl S, “Telephone Versus Face-To-Face Interviews: Mode Effect on Semi-Structured
583 Interviews With Children.” *Sociological Methodology*, vol. 43(1), pp. 133-177, 2013.
- 584 Wild D, Grove A, Martin M, et al., “ISPOR Task Force for Translation and Cultural Adaptation.
585 Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-
586 Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and
587 Cultural Adaptation.” *Value Health*, vol. 8(2), pp. 94-104, 2005.
- 588 Wilson H, Dashiell-Aje E, Anatchkova M, et al. “Beyond study participants: a framework for
589 engaging patients in the selection or development of clinical outcome assessments for
590 evaluating the benefits of treatment in medical product development.” *Quality of Life*
591 *Research*, Vol. 27(1), pp. 5-16, 2018.
- 592

593 **APPENDIX 1. Other Qualitative Methods**

594
595 In addition to one-on-one interviews and focus groups, there are other qualitative methods that
596 can be used to elicit what is important to patients, which are described in the following sections.
597

598 *1. Delphi methods*

599
600 The Delphi Panel technique is a multistage survey process with the intent to achieve
601 consensus among experts, including patients, on a topic or issue. It can provide valuable data
602 to help describe a patient experience or event.²³
603

604 *2. Observational methods*

605
606 Observational research methods can involve observations of patients by the researcher in a
607 naturalistic setting (e.g., home or school), a research facility, or virtual environment (e.g.,
608 online communities, social media) to generate data related to symptoms or daily life
609 functioning. These methods often involve assessment of events and patient attitudes and
610 behaviors over a period of time. Observational methods might be used to help understand
611 experiences described through other methods.
612

613 Examples of scenarios where these methods could be useful include (but are not limited to):
614

- 615 • In-person observations of children with attention deficit hyperactivity disorder in a
616 classroom setting
- 617
- 618 • Room surveillance that can be live or through use of video recordings to capture behavior
619 while sleeping
- 620
- 621 • Room surveillance for observation of aggressive behaviors or confusion in patients with
622 advanced Alzheimer’s disease
- 623
- 624 • Social media listening (e.g., observing interactions among social media users in an online
625 community) to understand how patients with a disease or condition describe their
626 experience with treatment
627

628 *3. Facilitated discussions at patient meetings*

629
630 Facilitated discussions in well-organized public meetings that include patients, caregivers
631 and patient representatives can generate useful public input and patient perspectives in
632 specific disease areas or topics. The Food and Drug Administration (FDA) has organized and
633 led such meetings under its PFDD initiative. FDA also welcomes patient organizations to

²³ There are many different Delphi methods described in the literature that can generate consensus data (Keeney et al., 2010).

Contains Nonbinding Recommendations

Draft — Not for Implementation

634 identify and organize patient-focused collaborations to generate public input on other disease
635 areas, using the process established by FDA-led PFDD meetings as a model.²⁴

636

637 4. *Survey instruments with open-ended questions*

638

639 See [section IV](#).

640

641

642 APPENDIX 2. Considerations for Selection of Qualitative Data Collection Methods

643

644 **Table 5. Some Potential Strengths and Limitations for Different Qualitative Data**
645 **Collection Methods**

| Data Collection Method | Strengths | Limitations |
|------------------------|--|---|
| One-on-one interviews | <ul style="list-style-type: none">• Can gain in-depth and broad information on the topic of interest, including nuanced data about an individual’s experience and perspectives• Can gain an understanding of how a respondent interprets a question that might be included in a questionnaire• Flexible format – can tailor interviews to generate appropriately detailed information based on research needs (e.g., through use of probing questions)• Greater scheduling flexibility compared with focus groups• Privacy and confidentiality – some people may be reluctant to share certain things in a group setting• Can be conducted in-person at a study site or at a person’s home (e.g. for severely ill patients), or via telephone or video conference | <ul style="list-style-type: none">• Duration (e.g., length of time it takes to conduct several patient interviews)• Participants may be uncomfortable providing complete or truthful information on sensitive topics to interviewers in person• Studies can be expensive (i.e., staff time for conducting multiple individual interviews) |

(Table continued)

²⁴ See <https://www.fda.gov/drugs/developmentapprovalprocess/ucm579400.htm>.

Contains Nonbinding Recommendations

Draft — Not for Implementation

| Data Collection Method | Strengths | Limitations |
|-------------------------------|---|--|
| Focus groups | <ul style="list-style-type: none"> • Can gain in-depth information on the topic of interest • Flexible format (see above one-on-one interviews) • Efficiency - Elicit feedback from multiple participants at one time • Participants can react to and build on each other’s ideas • Relatively inexpensive | <ul style="list-style-type: none"> • May not be efficient in covering maximum depth on an issue • Participants may become distracted by other participants in the group • Participants may experience peer-pressure within the group • Single individuals might dominate the conversation preventing multiple perspectives from being shared • Group setting may inhibit some individuals from providing sensitive information • Less flexibility in scheduling for a group of people which can lead to recruitment challenges |
| Delphi panels | <ul style="list-style-type: none"> • May provide a method for reaching consensus among appropriate experts and stakeholders on important issues and topics • Anonymous²⁵ process, when appropriate, reduces the role of ego and interpersonal issues in reaching consensus • Information can be collected remotely (e.g., via email or file sharing software) | <ul style="list-style-type: none"> • Lack of universal guidelines for process • Definitions of “expert” opinion are variable • No clear standards for acceptable level of consensus • Size of expert panel should be considered as it is difficult to achieve consensus among a larger group • Implications for lack of anonymity in the case of modified Delphi panel methods • Can be time-consuming and costly (e.g., high key opinion leader remuneration costs) |

(Table continued)

²⁵ Responses are anonymous only to group members; researchers are aware of respondent identities. Similar to methods used in reporting aggregated data for interviews and focus groups, responses will be reported using a unique identifier assigned to each expert.

Contains Nonbinding Recommendations

Draft — Not for Implementation

| Data Collection Method | Strengths | Limitations |
|---|---|---|
| <p>Observations of patient behavior or events (e.g., in real-world settings and in real-time; social media listening)</p> | <ul style="list-style-type: none"> • May be low burden for participants because the observation does not require active participation • Potential to observe episodic behavior and signs in a real-world context • Observations do not rely on patient or caregiver report | <ul style="list-style-type: none"> • May be time-consuming and logistically cumbersome to execute if conducted in natural settings (e.g., study environments may vary across locations) • Patient and others' privacy needs to be addressed given patients will be observed in their daily lives • Some concepts and experiences are not observable • Can be expensive (e.g., equipment if recording behaviors, staff time for observing in real-time) • Participant behavior may be affected by observer presence • If conducted in naturalistic settings, may be variable and affect the reliability and generalizability of the results • May call for observers to receive special training on identifying relevant observations (i.e., deciding what observations are important or unimportant) |
| <p>Facilitated discussions in organized patient conferences/meetings</p> | <ul style="list-style-type: none"> • Gain in-depth information on the topic of interest • Efficiency - Elicit feedback from multiple participants at one time • Can include real-time public polling exercises | <ul style="list-style-type: none"> • Input is limited to patients who can attend the meeting, which may affect the reliability and generalizability of the results • Although panelists speak to the moderator, participants do not interact with each other in the same way that focus group participants do • Representativeness and clinical confirmation of diagnosis may be difficult to determine. • See potential limitations of focus groups |

Contains Nonbinding Recommendations

Draft — Not for Implementation

647 **APPENDIX 3. Study Materials for Qualitative Studies**

648

649 [Table 6](#) discussed considerations of special relevance to designing and implementing study materials for qualitative studies of patient
650 experience.

651

652 **Table 6. Study Materials²⁶**

| Study Material | Components | Considerations |
|----------------------------|--|---|
| Study protocol | <ul style="list-style-type: none"> • Details on how the research will be conducted • Evidence to support the conduct of the study (e.g., unmet need) • Description of all research-related activities and study activities that patients will undergo | <ul style="list-style-type: none"> • Outline clear research objectives and questions • Specify details on target population, including demographics, clinical characteristics (e.g., phenotype, genotype, disease severity), and other pertinent characteristics (e.g., geographic representation) • Specify how data will be prepared for analysis (e.g., transcription, audio-/video-recorded, internet data, metadata, archives) • Include information regarding projected clinical site enrollment characteristics (e.g., geographic location; referral/academic centers versus community centers) to help further characterize the study sample • See Guidance 1 for details regarding considerations for study sampling and representativeness • Identify the number and duration of discussion sessions you plan to conduct; this should be dependent on: <ul style="list-style-type: none"> – Number objectives and research questions – Level of heterogeneity (e.g., age, <i>sex</i>, in the target population) – Number of subgroups (e.g., disease severity levels, phenotypes, informants [just patients or patients and their caregivers]) |
| Interview/discussion guide | <ul style="list-style-type: none"> • Interviewer/facilitator instructions • Study instruction • Warm-up questions • Core topic-related questions • Wrap-up questions • Discussion conclusion | <ul style="list-style-type: none"> • Use terms participants can understand and avoid technical terms where possible (e.g., choose to use the term “shortness of breath” rather than “dyspnea”). • Avoid asking leading questions that guide participants to respond with a preferred answer. • Avoid asking questions that imply you are casting judgment on a participant’s beliefs or choices. |

(Table continued)

²⁶ For the documents discussed in the table, there may also be other generally applicable regulations, guidance(s), standards and/or requirements.

Contains Nonbinding Recommendations

Draft — Not for Implementation

| Study Material | Components | Considerations |
|-----------------------|--|--|
| | | <ul style="list-style-type: none"> • Use open-ended questions rather than closed-ended questions, where appropriate, to elicit spontaneous information from participants. • Frame questions within the context of a participant’s experiences; avoid questions about abstract or theoretical concepts to the extent possible. • Consider eliciting specific data by framing questions using targeted approaches such as): <ul style="list-style-type: none"> – Diary questions (patients asked to describe a typical day) – Critical incidents (patient reports worst/best experience) – Free listing (patients list all symptoms, impacts, treatments, etc.) – Ranking (patients rank importance of symptom, treatment benefit, etc.) |
| Training Materials | <ul style="list-style-type: none"> • Detailed coverage of the protocol contents • Consent/assent forms • Mock discussion session (staff can evaluate flow of discussion) | <ul style="list-style-type: none"> • Train staff using standardized training materials (e.g., training documents, PowerPoint slides) • Provide refresher training |
| Glossary | <ul style="list-style-type: none"> • Definitions of terminology | <ul style="list-style-type: none"> • Clearly define key terminology within the qualitative text and ensure consistent terminology is used throughout study document(s) |
| Coding dictionary | <ul style="list-style-type: none"> • Codes (category or concept descriptions) • Coding structure • Memos (ideas or thoughts how code derived) | <ul style="list-style-type: none"> • Outline clear instructions for categorization, including code definitions, instructions, and considerations • Derive initial codes from prior knowledge (e.g., natural history, conceptual model, disease model, discussion guide structure) • Creating too many codes or nuanced categories may make it difficult for coders to capture and interpret concepts during the data analysis phase |
| Data analysis plan | <ul style="list-style-type: none"> • Analytic methods, including coding software • Identification of coders/analysts (including credentials) • Plans for resolving discrepancies among coders and other quality assurance | <ul style="list-style-type: none"> • Determine sample size needed for the study • Identify and specify appropriate analytic methods for data type • Consider what approach would be most appropriate to present data (tables, figures, etc.) |

(Table continued)

Contains Nonbinding Recommendations

Draft — Not for Implementation

| Study Material | Components | Considerations |
|-----------------------|--|-----------------------|
| | <p>measures (e.g., intra-rater reliability; Kappa statistic)</p> <ul style="list-style-type: none">• Description of coding stages (e.g., initial coding, interim checks – including plans for coding dictionary refinement)• Plans for data visualization• Table/figure shells | |

653

Contains Nonbinding Recommendations

Draft — Not for Implementation

654 APPENDIX 4. Analysis of Qualitative Data

655
656 Qualitative data can be voluminous, so it is important to have a standardized method to analyze
657 and interpret the volume of data in a practical and consistent way. Qualitative data should be
658 prepared before analysis. Preparation can include aggregation or transcription of data from
659 different sources, including:

- 660
- 661 • In-person interviews or focus groups
- 662
- 663 • Video-/online-recordings
- 664
- 665 • Internet (e.g., social media, chat room dialogues)
- 666
- 667 • Metadata (e.g., date of interview, name of interviewer, demographic details of
668 respondent, source of field notes, initial ideas of analysis)
- 669

670 [Table 7](#) provides considerations for analyzing qualitative data. Note the steps for data analysis in
671 qualitative studies may be iterative and are not necessarily sequential.

672

673 **Table 7. Steps Typically Used for Data Analysis in Qualitative Studies**

| Steps | Description |
|-----------------------------------|--|
| Compiling and organizing data | <ul style="list-style-type: none">• Arrange notes from research and other data collection in a useful and standardized order (e.g., electronic storage, computer programs) |
| Describing and classifying data | <ul style="list-style-type: none">• Break down compiled data into smaller pieces• Reorganize pieces into different groupings/sequences (e.g., codes) |
| Interpreting data | <ul style="list-style-type: none">• Use the grouped/sequenced data to identify the larger meaning of the data• Connect concepts from the data to other evidence (e.g., relevant literature, expert opinion)• Evaluate whether no new and important concepts have appeared (i.e., saturation) |
| Representing and Visualizing Data | <ul style="list-style-type: none">• Package data in a way that can be easily understood (e.g., text, tables, figures) |

674
675 Transcripts should be analyzed using methods appropriate for categorization and aggregation of
676 study results. There are different approaches to describe and classify qualitative data, some that
677 may involve coding and some that may not. The FDA generally recommends that qualitative
678 data are coded for regulatory submissions.

679
680 If a coding approach is selected for analysis, considerations commonly include but are not
681 limited to the following:

- 682
- 683 • Select the appropriate coding approach for the data of interest.
- 684
- 685 • Determine the appropriate level of detail for what is to be coded (e.g., line-by-line coding
686 or select segments of text).

Contains Nonbinding Recommendations

Draft — Not for Implementation

687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727

- Decide what data is relevant enough to be coded.
- Move methodically to a slightly higher conceptual level initially when coding data.
- Carefully consider the grammatical form of the coded words (e.g., actions versus processes versus nouns).
- Ensure codes are applied consistently to all data.
- Calculate interrater agreement among multiple coders.
- Apply quality assurance checks throughout the coding and analysis process.

When you have literature, expert input, and appropriate knowledge, a coding dictionary²⁷ is generally developed before coding begins, and standardized codes are used to categorize transcript data. In most instances, this standardized coding method is used along with a more emergent coding method where concepts not identified beforehand can be incorporated into the analyses. The coding dictionary would then evolve as new concepts are identified and emerge from the data. See examples below.

If a coding approach is not selected for analysis, methods that are commonly used include but are not limited to the following:

- Arrange notes (notes about original data) in a thematical manner.
- Ensure your notes precisely cite the original data (or precisely locate the places in the database).
- Implement a procedural check (take notes and crosswalk them backwards into the original database).

It is important to note that regardless of analytic method, you should maintain a methodical analytical procedure to avoid nonsystematic and inconsistent judgments.

²⁷ A coding dictionary is a guide with predetermined concept categories and descriptions (related to the research objectives and questions) that is developed before data collection and analysis.

Contains Nonbinding Recommendations

Draft — Not for Implementation

728

| | |
|--|--|
| EXAMPLE | |
| Coding line-by-line (applying codes to each line of qualitative data) | |
| <p>Fatigue Time-sensitive medication Interference with daily activities Limits physical functioning Rash Itchy</p> | <p>01 INTERVIEWER 02 How do you feel when you take your 03 medicine? 04 PATIENT 05 I feel extremely tired after taking my medicine. I 06 am not sure if it is related to the time of day that I 07 take it or not. Regardless, I cannot complete chores 08 around the house or take long walks. 09 I also have noticed a rash along my upper arm, 10 which has caused a lot of itching.</p> |
| EXAMPLE | |
| Using data to generate themes and codes | |
| <p>01 PATIENT 02 Because I was in extreme pain, my doctor wanted to 03 re-evaluate some of my meds. The doctor told me I 04 would have to stay in the hospital for monitoring. I 05 was afraid that this was would not be covered under 06 my insurance. I ended up calling my family to see if 07 they could visit me.</p> | <p>Pain Requesting regimen evaluation Hospitalization Medical access Family support</p> |

729

730 Qualitative data should be presented in a clear manner. It is generally helpful to include participant
 731 statements, in the participants’ own words, to represent the qualitative data. Stakeholders should use
 732 their best judgment on how best to present the data. [Table 8](#) describes three possible modes to
 733 display qualitative data.

734

735 **Table 8. Modes for Displaying Qualitative Data**

| Type of Display | Illustrative Example |
|-----------------------|---|
| Word tables and lists | Summary of findings placed in a table or matrix of rows and columns Chronology Summarize characteristics (e.g., demographics) of participants studied or interviewed List of de-identified individual participants in a study (usually using pseudonyms) and their study characteristics (other than demographics) |
| Graphics | Hierarchical chart (e.g., tree diagram, conceptual framework) Flowchart Spatial layout of a study area |
| Pictures | Photographs Reproductions (e.g., participant’s drawings or pictures) |

736

Contains Nonbinding Recommendations

Draft — Not for Implementation

737 To improve quality control of data, consider the following:

738

739 • Thoroughly read all qualitative transcripts and reread the transcripts.

740

741 • Double code qualitative data to ensure reliability.

742

743 • Create an audit trail.

744

745 • Concepts emerging from the interviews should be analyzed and summarized in sets in the
746 order that the data are collected (i.e., as interviews are conducted) and displayed in a
747 saturation table or grid. Although there are no set criteria for how saturation should be
748 undertaken, the steps indicated are representative of most approaches.

749

EXAMPLE

Concepts reported in the first 25 percent of planned interviews with patients are compared to the next 25 percent of planned interviews after they are conducted. Both sets of interviews (50 percent of the originally planned number) are compared with the next 25 percent of the planned interviews, and subsequently all these interviews (75 percent) are compared to the next 25 percent interviews and so on. The goal of the saturation process is to compare the amount of new information that is observed in the first interview set to the second interview set and so forth. Interviews are typically conducted until saturation is met, and no new concepts are emerging from the last round of interviews.

750

751 **Table 9** shows a saturation grid example summarizing focus group data. In this example, the
752 researchers identified two symptoms (i.e., Symptom A and Symptom B) based on literature
753 review and subject matter expert input. In addition, the researchers identified additional
754 symptoms based on the transcripts (i.e., Other Emergent Symptoms). Although no new emergent
755 concepts were identified at Site 4 it may also be useful (if feasible) to perform one or more
756 additional focus groups to confirm there are not additional emergent symptoms. Additionally,
757 before assuming that concept saturation may have been met, it is important to review focus group
758 participant demographics to assess representativeness.

759

Contains Nonbinding Recommendations

Draft — Not for Implementation

760 **Table 9. Concept Saturation Table Example (N=6 patients per focus group; 24 total**
761 **patients) ***

| Concept | Group 1 | Group 2 | Group 3 | Group 4 |
|-----------|---------|---------|---------|---------|
| Symptom A | 5 | 4 | 3 | 3 |
| Symptom B | 4 | 3 | 5 | 3 |
| Symptom C | 3 | 4 | 4 | 3 |
| Symptom D | | 2 | 1 | 2 |
| Symptom E | | 3 | 2 | |
| Symptom F | 4 | 3 | 5 | 4 |
| Symptom G | | | 1 | |

762 * Highlighted cells indicate the first time a concept is mentioned, and numbers within the
763 cells indicate the number of patients in that focus group endorsing the concept.

764
765 Regarding the reporting of qualitative research, the following literature references contain useful
766 suggestions:

- 767
- 768 • Tong A, Sainsbury P, Craig J, “Consolidated Criteria for Reporting Qualitative Research
769 (COREQ): A 32-item Checklist for Interviews and Focus Groups.” *International Journal*
770 *for Quality Health Care*, vol. 19(6), pp. 349-357, 2007.
 - 771 • Gibbs, GR., *Qualitative Research Kit: Analyzing Qualitative Data*, London, England:
772 SAGE Publications Ltd., 2007.
- 773
774

Contains Nonbinding Recommendations

Draft — Not for Implementation

775 **APPENDIX 5. Screening and Exit Interview Studies/ Survey Studies**

776

777 Screening/exit interviews and survey instruments may be implemented within the context of a
778 clinical trial. They can be helpful in obtaining patient feedback regarding various topics, such as
779 the following:

780

781 • Reported symptom changes (benefits, tolerability and unintended effects) experienced by
782 patients throughout a trial

783

784 • Participant treatment expectations

785

786 • Anticipated and unintended symptoms and adverse events

787

788 • Viability of proposed dosing regimen

789

790 • Patients' experience with clinical trial participation, e.g.:

791

792 ○ In blinded trials, whether they thought they could tell whether they were on the
793 experimental treatment (or not) and why they thought they were on that treatment

794

795 ○ Thoughts regarding study procedures

796

797 ○ Experience with modes of data collection (user experience with electronic data
798 entry)

799

800 • Benefit-risk perspective(s) from the patient/caregiver

801

802 The following are examples of potential strengths associated with conducting screening/exit
803 interviews and survey instruments:

804

805 • In rare diseases, they can contribute cumulative evidence on demographics, medical
806 history, and aspects of the patient experience.

807

808 • They can inform initial development or refinement of a clinical outcome assessment in
809 early medical product development through cognitive interviews as part of a mixed-
810 method approach.

811

812 • They can add greater depth to data in rare diseases (or possibly other diseases with not
813 much patient input) where stand-alone qualitative studies are less feasible.

814

815 • They can be used to obtain participant input on meaningful outcomes or meaningful
816 change by eliciting patient definitions of symptom improvement, stability, or worsening

817

Contains Nonbinding Recommendations

Draft — Not for Implementation

818 Potential limitations of screening/exit interviews and survey instruments include:

819

820 • Extra burden on site staff (e.g., additional operational procedures beyond the clinical
821 study)

822

823 • Extra burden for patients/caregivers, on top of standard clinical trial protocol

824

825 • For interviews, issues might arise regarding interview scheduling, administration time
826 and confidentiality (e.g., certain sites/countries cannot share participant contact details
827 with third-party vendors who might be conducting the interviews)

828

829 If screening/exit interviews are implemented, FDA generally recommends assessment of the
830 site's experience with varying levels of interview complexity before designing the study to help
831 determine who should conduct the study (e.g., site staff, vendor), with interview protocols and
832 interviewer guides developed thoughtfully, keeping in mind the context of the individual study
833 design. Likewise, interviews should generally be conducted before (i.e., screening interviews) or
834 after (i.e., exit interviews) patients complete the main portion of the clinical study to avoid any
835 potential compromise of trial integrity.

836

837 Future guidance documents will discuss other considerations for developing and administering
838 screening/exit interviews and survey instruments.

839 **APPENDIX 6. Glossary**

840

841 This glossary defines terms that will be used in the series of methodological Patient-Focused
842 Drug Development (PFDD) FDA guidance documents that are required by the 21st Century
843 Cures Act, and part of commitments made by FDA under the sixth authorization of the
844 Prescription Drug User Fee Act (PDUFA VI). The goal of this glossary is to provide
845 standardized nomenclature and terminologies related to patient-focused medical product
846 development. As appropriate, definitions from existing federal resources (e.g., Biomarkers,
847 EndpointS, and Other Tools (BEST) Resource)²⁸ have been incorporated into this glossary.
848 External resources were also used to define terms and are cited.

849

850 **1. Ability to Detect Change:** Evidence that a COA can identify differences in scores over time
851 in individuals or groups who have changed with respect to the measurement concept.

852 **2. Assent:** A child’s affirmative agreement to participate in research capture through verbal and
853 written acknowledgement. Mere failure to object should not, absent affirmative agreement,
854 be construed as assent.

855 **3. Benefit:** Benefits are the favorable effects of a medical product. Types of benefit include
856 clinical benefit (see *clinical benefit*). Benefits may also include important characteristics of
857 the medical product, such as convenience (e.g., a more convenient dosing regimen or route of
858 administration) that may lead to improved patient compliance, or benefits that affect those
859 other than the patient. (Source: [International Conference on Harmonisation \(ICH\) Guideline,
860 Revision of M4E Guideline on Enhancing the Format and Structure of Benefit-Risk
861 Information in ICH \(Efficacy – M4E\(R2\)\), available at
862 \[http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4E_R2_Efficacy/M4E
864 \\[R2_Step_4.pdf\\]\\(#\\); ANSI/AAMI/ ISO 14971: 2007/\\(R\\)2016 Medical devices—Application of
risk management to medical devices.\]\(http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4E_R2_Efficacy/M4E
863 _R2_Step_4.pdf\)\)](#)

865 **4. Benefit-Risk Assessment:** Evaluation of the demonstrated benefits and risks of a medical
866 product and making a judgment as to whether the expected benefits outweigh the potential
867 risks associated with its expected use.

868 **5. Disease burden:** The impacts, direct and indirect, of the patient’s health condition that have
869 a negative effect on his or her health, functioning, and overall well-being. Disease burden
870 includes but is not limited to the physical and physiologic impacts of the disease and its
871 symptoms; co-morbidities; emotional and psychological effects of the disease, its
872 management, or its prognosis; social impacts; effects on relationships; impacts on the
873 patient’s ability to care for self and others; time and financial impacts of the disease and its
874 management; and considerations of the impacts on the patient’s family.

875 **6. Caregiver:** A person who helps a patient with daily activities, health care, or any other
876 activities that the patient is unable to perform because of illness or disability, and who

²⁸[Available at https://www.ncbi.nlm.nih.gov/books/NBK338448/](https://www.ncbi.nlm.nih.gov/books/NBK338448/)

877 understands the patient’s health-related needs. This person may or may not have decision-
878 making authority for the patient and is not the patient’s health care provider.

879 **7. Ceiling Effect:** A ceiling effect can occur at the item level or at the scale score level. An
880 item level ceiling effect is observed when a large concentration of participants endorses the
881 highest response category within an item. A scale score level ceiling effect is observed when
882 a large concentration of participants’ scores fall at or near the upper limit of the scale score of
883 the instrument. Either situation may occur when the upper extreme of the concept(s) assessed
884 by item response categories or by the scale score of the instrument does not sufficiently
885 match the level of the upper extreme of the target patient population.

886 **8. Clinical Benefit:** A positive clinically meaningful effect of an intervention (i.e., a positive
887 effect on how an individual feels, functions, or survives). (Source: [“BEST \(Biomarkers,
888 EndpointS, and other Tools\) Resource”](#))

889 **9. Clinical Outcome:** An outcome that describes or reflects how an individual feels, functions
890 or survives. (Source: [“BEST \(Biomarkers, EndpointS, and other Tools\) Resource”](#))

891 **10. Clinical Outcome Assessment (COA):** Assessment of a clinical outcome can be made
892 through report by a clinician, a patient, a nonclinician observer or through a performance-
893 based assessment. Types of COAs include: patient-reported outcome (PRO) measures,
894 clinician-reported outcome (ClinRO) measures, observer-reported outcome (ObsRO)
895 measures, and performance outcome (PerfO) measures. (Source: [“BEST \(Biomarkers,
896 EndpointS, and other Tools\) Resource”](#))

897 **11. Clinician-Reported Outcome (ClinRO):** A measurement based on a report that comes from
898 a trained health-care professional after observation of a patient’s health condition. Most
899 ClinRO measures involve a clinical judgment or interpretation of the observable signs,
900 behaviors, or other manifestations related to a disease or condition. ClinRO measures cannot
901 directly assess symptoms that are known only to the patient (e.g., pain intensity). (Source:
902 [“BEST \(Biomarkers, EndpointS, and other Tools\) Resource”](#))

903 **12. Cognitive Interviewing:** A qualitative research process used to determine whether concepts
904 and items are understood by respondents in the same way that instrument developers intend.
905 Cognitive interviews involve incorporating follow-up questions in a field test interview to
906 gain a better understanding of how respondents interpret questions/tasks asked of them. In
907 this method, respondents are often asked to think aloud and describe their thought processes
908 as they answer the instrument questions. Respondents should reflect the target population
909 who will be responding to the instrument during the study.

910 **13. Concept (also referred to as concept of interest):** In a regulatory context, the concept is the
911 aspect of an individual’s clinical, biological, physical, or functional state, or experience that
912 the assessment is intended to capture (or reflect). (Source: [“BEST \(Biomarkers, EndpointS,
913 and other Tools\) Resource”](#))

- 914 **14. Concept Elicitation:** A process or method to collect a holistic set of relevant concepts (i.e.,
915 disease and treatment symptoms and associated impacts) that are important to patients from
916 relevant stakeholders (e.g., patients, experts, caregivers).
- 917 **15. Concept Saturation:** When interviewing participants representative of the target patient
918 population, concept saturation is the point at which no new important concepts relevant to the
919 research question are emerging from iterative rounds of interviews; collecting additional data
920 will not likely add to the understanding of how participants perceive the concept of interest.
- 921 **16. Conceptual Framework:** An explicit description or a diagram for an instrument showing the
922 relationships between items (i.e., questions/tasks included in the instrument), domains (sub-
923 concepts), and concepts measured, and the scores produced by a COA. The conceptual
924 framework of a COA evolves over the course of instrument development as empiric evidence
925 is gathered to support item grouping and scores.
- 926 **17. Construct Validity:** Evidence that relationships among items, domains, and concepts
927 conform to a priori hypotheses concerning logical relationships that should exist with other
928 measures or characteristics of patients and patient groups.
- 929 **18. Content Validity:** Evidence from qualitative research demonstrating that an instrument
930 measures the concept of interest, including evidence that the items and domains of an
931 instrument are appropriate and comprehensive relative to its intended measurement concept,
932 population, and use. Testing other measurement properties will not replace or rectify
933 problems with content validity.
- 934 **19. Context of Use:** A statement that fully and clearly describes the way a medical product
935 development tool is to be used and the medical product development-related purpose of the
936 use. (Source: [“BEST \(Biomarkers, EndpointS, and other Tools\) Resource”](#))
- 937 **20. Criterion Validity:** The extent to which the scores of a COA are related to a known gold
938 standard measure of the same concept. For most COAs, criterion validity cannot be measured
939 because there is no gold standard.
- 940 **21. Data Analysis Plan:** A roadmap for how data will be organized and analyzed and how
941 results will be presented. A data analysis plan should be established when planning a
942 research study (i.e., before data collection begins). Among other things, the data analysis plan
943 should describe: (a) the data to be collected; (b) the analyses to be conducted to address the
944 research objectives, including assumptions required by said analyses; (c) data cleaning and
945 management procedures; (d) data transformations, if applicable; and (e) how the study results
946 will be presented (e.g., graphs, tables).
- 947 **22. Data Management Plan (DMP):** A written document that describes the data you expect to
948 acquire or generate during the course of your research study; how you intend to manage,
949 describe, analyze, and store said data; and what mechanisms you will use at the end of your
950 study to preserve and share your data. (Source: [Stanford University Libraries n.d.\(b\), “About
951 Data Management Plans \(DMPS\),” available at https://library.stanford.edu/research/data-
952 management-services/data-management-plans.](#))

Contains Binding Provisions

Draft — Not for Implementation

- 953 **23. Domain:** A sub-concept represented by a score of an instrument that measures a larger
954 concept comprised of multiple domains. For example, psychological function is the larger
955 concept containing the domains subdivided into items describing emotional function and
956 cognitive function.
- 957 **24. Endpoint:** A precisely defined variable intended to reflect an outcome of interest that is
958 statistically analyzed to address a particular research question. A precise definition of an
959 endpoint typically specifies the type of assessments made; the timing of those assessments;
960 the assessment tools used; and possibly other details, as applicable, such as how multiple
961 assessments within an individual are to be combined. (Source: [“BEST \(Biomarkers,
962 EndpointS, and other Tools\) Resource”](#))
- 963 **25. Fit-for-Purpose:** A conclusion that the level of validation associated with a tool is sufficient
964 to support its context of use. (Source: [“BEST \(Biomarkers, EndpointS, and other Tools\)
965 Resource”](#))
- 966 **26. Floor Effect:** A floor effect can occur at the item level or at the scale score level. An item
967 level floor effect is observed when a large concentration of participants endorses the lowest
968 response category within an item. A scale score level floor effect is observed when a large
969 concentration of participants’ scores fall at or near the lower limit of the scale score of the
970 instrument. Either situation may occur when the lower extreme of the concept(s) assessed by
971 item response categories or by the scale score of the instrument does not sufficiently match
972 the level of the lower extreme of the target patient population.
- 973 **27. Focus Group:** A carefully planned discussion conducted among a small group of
974 participants, led by a moderator with the appropriate training. Group dynamics in a focus
975 group can facilitate additional insights that one-on-one interviews cannot.
- 976 **28. Generalizability:** The extent to which study findings can be reliably extended to the target
977 population of interest.
- 978 **29. Health Literacy:** The degree to which individuals have the capacity to obtain, process, and
979 understand basic health information and services needed to make appropriate health
980 decisions. (Source: U.S. Department of Health and Human Services, 2000) Health literacy
981 also includes numeracy skills—such as calculating cholesterol and blood sugar levels,
982 measuring medication doses, and understanding nutrition labels—and knowledge of health
983 topics.
- 984 **30. Instrument or Tool:** An assessment system comprising three essential components: (1)
985 materials for measurement; (2) an assay for obtaining the measurement; and (3) method
986 and/or criteria for interpreting those measurements. (Source: [“BEST \(Biomarkers,
987 EndpointS, and other Tools\) Resource”](#))
- 988 **31. Item:** An individual question, statement, or task (and its standardized response options) that
989 is evaluated or performed by the patient to address a particular concept.

Contains Binding Provisions

Draft — Not for Implementation

- 990 **32. Item Tracking Matrix:** A record of the development (e.g., additions, deletions,
991 modifications, and reasons for the changes) of items or tasks used in an instrument.
- 992 **33. Literacy:** A person's ability to read, write, speak, and compute and solve problems at levels
993 necessary to (a) function on the job and in society; (b) achieve one's goals; and (c) develop
994 one's knowledge and potential. (Source: Public Law 102-73. The National Literacy Act of
995 1991)
- 996 **34. Measurement Properties:** All the attributes relevant to the application of a COA including
997 the content validity, construct validity, reliability, and ability to detect change. These
998 attributes are specific to the measurement application and cannot be assumed to be relevant
999 to all measurement situations, purposes, populations, or settings in which the instrument is
1000 used.
- 1001 **35. Methodologically Sound:** Assurance that the methods and processes used to obtain and
1002 analyze patient experience data are rigorous, robust, and adhere to scientifically established
1003 principles and best practices for method development or implementation. Evidence generated
1004 by methodologically sound methods and processes increases confidence that the results can
1005 be trusted, interpreted, and support the intended regulatory uses.
- 1006 **36. Mixed-Method Research:** Research that uses both qualitative and quantitative research
1007 methods. See *qualitative and quantitative research methods*.
- 1008 **37. Observer-Reported Outcome (ObsRO):** A measurement based on a report of observable
1009 signs, events, or behaviors related to a patient's health condition by someone other than that
1010 patient or a health professional. Generally, ObsROs are reported by a parent, caregiver, or
1011 someone who observes the patient in daily life, and ObsROs are particularly useful for
1012 patients who cannot report for themselves (e.g., infants or individuals who are cognitively
1013 impaired). An ObsRO measure does not include medical judgment or interpretation. (Source:
1014 [“BEST \(Biomarkers, EndpointS, and other Tools\) Resource”](#))
- 1015 **38. Observational Research:** A type of nonexperimental social science research technique in
1016 which a researcher directly observes ongoing phenomena in a natural setting. In health
1017 sciences, this can include, but is not limited to, observing behaviors and disease signs
1018 (tremors) in real-world settings and in real-time.
- 1019 **39. Patient:** Any individual with or at risk of a specific health condition, whether the individual
1020 currently receives any therapy to prevent or treat that condition. Patients are the individuals
1021 who directly experience the benefits and harms associated with medical products.
- 1022 **40. Patient Advocate:** An individual or group of individuals who may or may not be part of the
1023 target patient population and who has a role in promoting an interest or cause to influence
1024 policy with respect to patients' health or health care.
- 1025 **41. Patient-Centered:** See *patient-focused*.

Contains Binding Provisions

Draft — Not for Implementation

- 1026 **42. Patient-Centered Outcome:** An outcome that is important to patients’ survival, functioning,
1027 or feelings as identified or affirmed by patients themselves, or judged to be in patients’ best
1028 interest by providers and/or caregivers when patients cannot report for themselves. (Source:
1029 ISPOR Plenary, [Patrick, 2013](#))
- 1030 **43. Patient Engagement:** Activities that involve patient stakeholders sharing their experiences,
1031 perspectives, needs, and priorities that help inform FDA’s public health mission. Such
1032 activities may include but are not limited to testimony at Advisory Committee meetings,
1033 submission to regulations.gov public docket; meetings attended by patients, FDA, and other
1034 stakeholders; other correspondence with FDA; interactions through social media; and
1035 interactions with or information from patient representatives or patient advocates.
- 1036 **44. Patient Experience Data:** Defined in Title III, section 3001 of the 21st Century Cures Act of
1037 2016, as amended by section 605 of the Food and Drug Administration Reauthorization Act
1038 (FDARA) of 2017,²⁹ and includes data that are collected by any persons and are intended to
1039 provide information about patients’ experiences with a disease or condition. Patient
1040 experience data can be interpreted as information that captures patients’ experiences,
1041 perspectives, needs, and priorities related to but not limited to (1) symptoms of their
1042 condition and its natural history; (2) impact of the conditions on their functioning and quality
1043 of life; (3) their experience with treatments; (4) input on which outcomes are important to
1044 them; (5) their preferences for outcomes and treatments; and (6) the relative importance of
1045 any issue as defined by patients.
- 1046 **45. Patient-Focused** (also referred to as *patient-centered*): Ensuring that patients’ experiences,
1047 perspectives, needs, and priorities are meaningfully incorporated into decisions and activities
1048 related to their health and well-being.
- 1049 **46. Patient-Focused Drug Development (PFDD)** (also referred to as *patient-focused medical*
1050 *product development*): A systematic approach to help ensure that patients’ experiences,
1051 perspectives, needs, and priorities are captured and meaningfully incorporated into the
1052 development and evaluation of medical products throughout the medical product life cycle.
- 1053 **47. Patient Input:** Information that captures patients’ experiences, perspectives, needs, and
1054 priorities. See also *patient experience data*
- 1055 **48. Patient Partner:** An individual patient, caregiver, or patient advocacy group that engages
1056 other stakeholders to ensure the patients’ wants, needs, and preferences are represented in

²⁹ “PATIENT EXPERIENCE DATA.—For purposes of this section, the term ‘patient experience data’ includes data that (1) are collected by any persons (including patients, family members and caregivers of patients, patient advocacy organizations, disease research foundations, researchers, and drug manufacturers); and (2) are intended to provide information about patients’ experiences with a disease or condition, including (A) the impact (including physical and psychosocial impacts) of such disease or condition, or a related therapy, on patients’ lives; and (B) patient preferences with respect to treatment of such disease or condition.” This definition is found in section 569C(c) of the FD&C Act (codified at 21 U.S.C. 360bbb–8c) and is referred to in section 3002 of the 21st Century Cures Act, which directed FDA to issue certain guidance documents regarding the collection of patient experience data (see section 3002(b)).

Contains Binding Provisions

Draft — Not for Implementation

- 1057 activities related to medical product development and evaluation. (Source: Wilson et al,
1058 2018)
- 1059 **49. Patient Perspective:** A type of patient experience data that specifically relates to patients’
1060 attitudes or points of view about their condition or management of their condition. Patient
1061 perspectives may include, but are not limited to, perceptions, goals, priorities, concerns,
1062 opinions, and preferences.
- 1063 **50. Patient Preference:** A statement of the relative desirability or acceptability to patients of
1064 specified alternatives or choice among outcomes or other attributes that differ among
1065 alternative health interventions. (Source: [FDA Guidance for Industry: Patient Preference
1066 Information –Voluntary Submission, Review in Premarket Approval Applications,
1067 Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in
1068 Decision Summaries and Device Labeling](#))
- 1069 **51. Patient-Reported Outcome (PRO):** A measurement based on a report that comes directly
1070 from the patient (i.e., study subject) about the status of a patient's health condition without
1071 amendment or interpretation of the patient's response by a clinician or anyone else. A PRO
1072 can be measured by self-report or by interview, provided that the interviewer records only the
1073 patient's response. Symptoms or other unobservable concepts known only to the patient (e.g.,
1074 pain severity or nausea) can only be measured by PRO measures. PROs can also assess the
1075 patient perspective on functioning or activities that may also be observable by others.
1076 (Source: [“BEST \(Biomarkers, EndpointS, and other Tools\) Resource”](#))
- 1077 **52. Patient Representative:** An individual, who may or may not be part of the target population,
1078 who has direct experience with a disease or condition (e.g., a patient or caregiver) and can
1079 provide information about a patient’s experience with the disease or condition.
- 1080 **53. Performance Outcome (PerfO):** A measurement based on a standardized task performed by
1081 a patient that is administered and evaluated by an appropriately trained individual or is
1082 independently completed.
- 1083 **54. Qualitative Research Methods:** Methods associated with the gathering, analysis,
1084 interpretation, and presentation of narrative information (e.g., spoken or written accounts of
1085 experiences, observations, and events). Qualitative research methods may also include direct
1086 observations (e.g., nonverbal communication and behaviors).
- 1087 **55. Quantitative Research Methods:** Methods associated with the gathering, analysis,
1088 interpretation, and presentation of numerical information.
- 1089 **56. Recall Period:** The period of time patients, caregivers, or clinicians are asked to consider in
1090 responding to a COA item or task. Recall can be momentary (real time) or retrospective of
1091 varying lengths.
- 1092 **57. Reliability:** The ability of a COA to yield consistent, reproducible estimates.

Contains Binding Provisions

Draft — Not for Implementation

- 1093 **58. Representativeness:** Confidence that a sample from which evidence is generated is
1094 sufficiently similar to the intended population. In the context of patient experience data,
1095 representativeness includes the extent to which the elicited experiences, perspectives, needs,
1096 and priorities of the sample are sufficiently similar to those of the intended patient
1097 population.
- 1098 **59. Research Protocol:** A document that describes the background, rationale, objectives, design,
1099 methodology, statistical considerations, and organization of a clinical research project.
1100 (Source: University of California San Francisco, 2017) A research protocol guides the study
1101 and associated data collection and analysis in a productive and standardized manner.
- 1102 **60. Response Scale:** The system of numbers or verbal anchors by which a value or score is
1103 derived for an item. Examples include verbal rating scale (VRS), numeric rating scale (NRS),
1104 and visual analog scale (VAS).
- 1105 **61. Risk Tolerance:** The degree to which a patient would accept increased probability or
1106 severity of a harm in exchange for a specific expected benefit. (Source: Medical Device
1107 Innovation Consortium, 2015)
- 1108 **62. Science of Patient Input:** Methods and approaches of systematically obtaining, analyzing,
1109 and using information that captures patients' experiences, perspectives, needs, and priorities
1110 in support of the development and evaluation of medical products.
- 1111 **63. Score:** A number derived from a patient's, caregiver's, or clinician's response to items or
1112 tasks in an instrument. A score is computed based on a prespecified, appropriate scoring
1113 algorithm and is subsequently used in statistical analyses of clinical trial results. Scores can
1114 be computed for individual items, domains, or concepts, or as a summary of items, domains,
1115 or concepts.
- 1116 **64. Scoring Algorithm:** A set of prespecified rules to assign numerical value or values to
1117 quantify the responses to the instrument. A scoring algorithm may create a single score from
1118 a single item or multiple items (e.g., domain score).
- 1119 **65. Sex:** The classification of living things, generally as male or female according to their
1120 reproductive organs and functions assigned by chromosomal complement. (Source: Institute
1121 of Medicine (US) Committee on Understanding the Biology of Sex and Gender Differences,
1122 2001)
- 1123 **66. Sign:** Any observable evidence of a disease, health condition, or treatment-related effect.
1124 Signs are usually observed and interpreted by the clinician but may be noticed and reported
1125 by the patient.
- 1126 **67. Social Media:** Web-based tools that are used for electronic communication. Social media
1127 may include but is not limited to (1) blogs; (2) microblogs; (3) social networking sites; (4)
1128 professional networking sites; (5) thematic networking sites; (6) wikis, (7) mashups; (8)
1129 collaborative filtering sites; (9) media sharing sites, and others. (Source: Grajales III et al.
1130 2014)

Contains Binding Provisions

Draft — Not for Implementation

- 1131 **68. Subgroup:** A subset of the study population or study sample defined by specific baseline
1132 characteristics. For example, demographic subgroups are commonly defined by subject sex,
1133 race, and age.
- 1134 **69. Symptom:** Any experience of a disease, health condition, or treatment-related effect that can
1135 be known and confirmed only by the patient, and therefore is most reliably assessed by direct
1136 patient report.
- 1137 **70. Target Population** (also referred to as *target patient population*, *underlying population*, or
1138 *intended population*): The group of individuals (patients) about whom one wishes to make an
1139 inference.
- 1140 **71. Target Product Profile (TPP):** A clinical development program summary in the context of
1141 labeling goals where specific types of evidence (e.g., clinical trials or other sources of data)
1142 are linked to the targeted labeling claims or concepts.
- 1143 **72. Task:** See *item*
- 1144 **73. Treatment burden:** The impacts of a specific treatment or treatment regimen that have a
1145 negative impact on a patient’s health, functioning, or overall well-being. Treatment burden
1146 includes but is not limited to side effects, discomfort, uncertainty about treatment outcomes,
1147 dosing and route of administration, requirements, and financial impacts.
- 1148 **74. Treatment Effect:** The amount of change in a disease/condition, symptom, or function that
1149 results from a medical intervention (as compared to not receiving the intervention or
1150 receiving a different intervention).
- 1151 **75. Treatment Outcome:** The benefits or harms to a patient who receives an intervention; the
1152 impact on a patient’s health, function, or well-being—or on a clinical indicator thereof—that
1153 is assumed to result from an intervention.
- 1154 **76. Usability Testing:** A formal evaluation with documentation of respondents’ abilities to use
1155 the instrument, as well as comprehend, retain, and accurately follow instructions.
- 1156 **77. User Acceptance Testing (UAT):** One aspect of an extensive system/software validation
1157 process designed to determine whether the software complies with the written system
1158 specification or user requirements document. It is not intended solely to determine whether
1159 respondents like or can use the system.
- 1160 **78. Validation:** A process to establish that the performance of a test, tool, or instrument is
1161 acceptable for its intended purpose. Elements of validation include but are not limited to the
1162 following: construct validation, content validation, criterion validation, analytical validation,
1163 clinical validation.