Patient Education and Counseling 99 (2016) 583–590

Contents lists available at ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pateducou

Communicating efficacy information based on composite scores in direct-to-consumer prescription drug advertising

Pamela A. Williams[a,*], Amie C. O'Donoghue[b], Helen W. Sullivan[b], Jessica Fitts Willoughby[a,c], Claudia Squire[a], Sarah Parvanta[a], Kevin R. Betts[b]

[a] RTI International, Research Triangle Park, NC, USA
[b] Office of Prescription Drug Promotion, Center for Drug Evaluation and Research, Food and Drug Administration (FDA), Silver Spring, MD, USA
[c] Edward R. Murrow College of Communication, Washington State University, Pullman, WA, USA

ARTICLE INFO

Article history:
Received 27 March 2015
Received in revised form 2 September 2015
Accepted 28 October 2015

Keywords:
United States
Direct-to-consumer (DTC) advertisements
Prescription drugs
Marketing
Communication
Composite scores

ABSTRACT

Objective: Drug efficacy can be measured by composite scores, which consist of two or more symptoms or other clinical components of a disease. We evaluated how individuals interpret composite scores in direct-to-consumer (DTC) prescription drug advertising.
Methods: We conducted an experimental study of seasonal allergy sufferers ($n = 1967$) who viewed a fictitious print DTC ad that varied by the type of information featured (general indication, list of symptoms, or definition of composite scores) and the presence or absence of an educational intervention about composite scores. We measured composite score recognition and comprehension, and perceived drug efficacy and risk.
Results: Ads that featured either (1) the composite score definition alone or (2) the list of symptoms or general indication information along with the educational intervention improved composite score comprehension. Ads that included the composite score definition or the educational intervention led to lower confidence in the drug's benefits. The composite score definition improved composite score recognition and lowered drug risk perceptions.
Conclusion: Adding composite score information to DTC print ads may improve individuals' comprehension of composite scores and affect their perceptions of the drug.
Practice implications: Providing composite score information may lead to more informed patient-provider prescription drug decisions.

1. Introduction

Consumers in the United States spent $320 billion on prescription drugs in 2011 [1], and spending is projected to increase [2]. Pharmaceutical companies frequently use direct-to-consumer (DTC) advertising to market their products to consumers [3]. In fact, 37% of the pharmaceutical industry's drug promotion dollars were spent on DTC advertising in 2011 [1]. DTC advertising has the potential to educate consumers about prescription drugs and contribute to informed decision-making [4–6]. However, it also has the potential to confuse consumers, leading them to overestimate drug benefits and underestimate drug risks [7–10].

The U.S. Food and Drug Administration (FDA) regulates DTC advertising practices [11]. For drugs to be approved, pharmaceutical companies must objectively demonstrate to FDA "substantial evidence" for the efficacy of their drugs, typically through two well-controlled clinical trials [12,13]. In some cases, drug efficacy can be measured by a single endpoint, such as high blood pressure [14]. Often, however, efficacy is measured by multiple endpoints that are sometimes combined into an overall score called a composite score, or a composite endpoint [15,16]. For example, nasal congestion is measured by examining individual symptoms such as runny nose, itchiness, and sneezing. While each symptom is measured on its own, an overall – or composite – score is computed from the individual symptom measurements. If a drug has a significantly better composite score than its comparison (i.e., treatment group vs. comparison group) with regard to the

* Corresponding author at: Social Policy, Health and Economics Research Unit, RTI International, 3040 East Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709-2194, USA. Fax: +1 919 541 7384.
E-mail addresses: pamwilliams@rti.org (P.A. Williams), jessica.willoughby@wsu.edu (J.F. Willoughby).

http://dx.doi.org/10.1016/j.pec.2015.10.019
0738-3991/

treatment of nasal congestion, it can be marketed for that health condition. However, it is possible for a drug to have a significantly better composite score even if it did not have a significantly better score on a particular symptom (e.g., runny nose). While scientists and medical professionals have had training to understand the difference between composite score endpoints and single endpoints, it is unclear whether consumers understand the difference.

There is a dearth of research examining consumers' understanding of composite scores in DTC ads; thus, O'Donoghue et al. [17] conducted a focus group study to explore how consumers react to DTC advertisements that present efficacy claims based on composite scores. Most participants had difficulty correctly interpreting efficacy information based on composite scores. After the focus group moderator explained the meaning of composite scores, some participants revised their opinion of the drug's efficacy and reported that the efficacy information was "much less convincing," often because it was unclear whether the drug would work for a particular symptom. As a result, some participants said they would want a drug ad to include information on the drug's efficacy for each component of the composite score. However, others felt that the ads already provided enough information and that adding more statistical details would make the ads more complicated, thus decreasing the likelihood that consumers would read them. The moderator used a decathlon as an example of a composite score and found that it resonated with participants. These focus group findings suggest that it may be worthwhile to examine further how including increasingly detailed information affects consumers' recognition that a drug's effectiveness is based on a composite score, comprehension of the concept of composite scores, and perceptions of a drug's efficacy and risk.

Given the paucity of research on consumer understanding of composite scores and how to best present this information in DTC advertisements, the overarching goal of the current research was to examine how DTC advertisements can effectively deliver composite endpoint efficacy information to maximize consumer comprehension and informed decision-making. The purpose of our study was to evaluate experimentally how consumers interpret and respond to DTC prescription drug advertising that includes efficacy information based on composite scores. We examined the effects of exposure to DTC advertisements that varied on

- the type of information that was presented (general indication information, a list of symptoms, or a definition of composite scores) and
- whether the information was accompanied by a brief educational intervention on composite scores.

We selected our manipulations of information type based on the way DTC ads conveyed this information at the time of the study. Most ads contained only a general indication statement, whereas some contained a list of components of composite scores, such as the symptoms of major depression. We contrasted these two conditions with a new condition we did not see in any DTC ads. In this condition we explicitly mentioned and defined composite scores. The brief educational intervention used the decathlon example from the focus groups to provide a transfer of the information to a real-life situation.

The following questions guided our analysis: do participants who experience an educational intervention differ on outcomes from participants who do not experience the educational intervention? Do participants who view the composite score definition differ on outcomes from participants who view the list of symptoms or general indication information? Is there an interaction effect between viewing the educational intervention and the type of information on the outcomes?

2. Methods

2.1. Sample

The study's sampling frame used GfK's (formerly Knowledge Networks) Internet panel, which is designed to be nationally representative and have fewer biases than opt-in Internet panels [18,19]. Adults who self-reported on GfK's profile panel survey that they had seasonal allergies were recruited through e-mail invitations sent to a random sample of this subset of panel members. Reminder invitations and a telephone reminder call were used to convert nonresponders. When panelists were still unresponsive, additional panelists were randomly sampled to replace the nonresponders. The final sample included 1967 participants. See Table 1 for demographic information.

2.2. Design

The study tested six manipulations of a fictitious drug ad. The experiment used a 3 (type of ad information: general indication, list of symptoms, or composite score definition) × 2 (educational intervention in ad: present or absent) between-subjects design.

Table 1
Summary of demographic characteristics of completed participants ($n = 1967$).

Characteristic	*n* (%)
Reported seasonal allergies	1967 (100%)
Sex	
Male	726 (36.9%)
Female	1241 (63.1%)
Age	
18–24	83 (4.2%)
25–34	232 (11.8%)
35–44	256 (13%)
45–54	426 (21.7%)
55–64	514 (26.1%)
65–74	333 (16.9%)
75+	123 (6.3%)
Race/ethnicity	
White	1705 (86.7%)
Black	127 (6.5%)
Other	134 (6.8%)
Ethnicity	
Hispanic or Latino	184 (9.4%)
Non-Hispanic	1783 (90.6%)
Education	
Less than high school diploma	47 (2.4%)
High school diploma or equivalent	481 (24.5%)
Some college	625 (31.8%)
Bachelor's degree or higher	814 (41.4%)
Household income	
Less than $20,000	232 (11.8%)
$20,000–29,999	168 (8.5%)
$30,000–39,999	184 (9.4%)
$40,000–49,999	195 (9.9%)
$50,000–74,999	390 (19.8%)
$75,000–99,999	339 (17.2%)
$100,000+	459 (23.3%)

Note: demographic data are unweighted. Percentages are based on non-missing data. One participant refused to answer items about race or ethnicity.

Fig. 1. Sample fictitious drug print advertisement for the composite score definition × education-absent condition.

2.3. Print stimuli

A realistic print ad for a fictitious seasonal allergy prescription drug, Trinase, was created for this experiment (see Fig. 1). The text of the ad was manipulated based on the type of information and the presence or absence of an educational intervention (see Table 2).

2.4. Procedure

All invited participants were randomly assigned to one of six experimental conditions. Participants could navigate back and forth between the pages of the ad without time limitations. After viewing the ads, participants completed a brief questionnaire that had been cognitively tested and pretested. Participants received points for participating in the study, which could later be exchanged for vouchers and gifts at a partner network. This study was approved by FDA's Research Involving Human Subjects Committee and RTI's Institutional Review Board.

2.5. Outcome measures

No prior measures of composite score understanding existed, so we drafted the questionnaire items based on our insights from the focus groups, cognitively tested them with 9 individuals who self-reported having seasonal allergies, and revised them based on testing results. We also were interested in the effects of the ads on the informed decision-making process, so we measured perceptions of drug efficacy and risk by adapting and fine-tuning questionnaire items that have been used in prior studies testing the effects of DTC ads [20–22].

*Composite score recognition*was measured with one "true" or "false" item that measured recognition of drug effectiveness as a basis for composite scores: "The ad stated that Trinase's

Table 2
Experimental conditions and stimuli manipulations.

Condition	General indication	List of symptoms	Composite score definition
Education absent	Trinase treats and helps prevent seasonal nasal allergy symptoms	Trinase treats and helps prevent seasonal nasal allergy symptoms: congestion, runny nose, nasal stuffiness, nasal itching, and sneezing	Trinase treats and helps prevent seasonal nasal allergy symptoms. Trinase's effectiveness is based on a composite score. A composite score is a single measure of how well a drug works based on a combination of symptoms. Trinase may not be as effective in addressing each symptom individually
Education present[a]	Trinase's effectiveness is based on a composite score. A composite score is like a decathlon. In that event, athletes compete in 10 events, such as the long jump, the shot put, and the 50-yard dash. An athlete may not win all events, but if he or she wins some and performs well enough in others, he or she may be the winner based on a combination of scores for each event	Trinase's effectiveness is based on a composite score. A composite score is like a decathlon. In that event, athletes compete in 10 events, such as the long jump, the shot put, and the 50-yard dash. An athlete may not win all events, but if he or she wins some and performs well enough in others, he or she may be the winner based on a combination of scores for each event	Trinase's effectiveness is based on a composite score. A composite score is like a decathlon. In that event, athletes compete in 10 events, such as the long jump, the shot put, and the 50-yard dash. An athlete may not win all events, but if he or she wins some and performs well enough in others, he or she may be the winner based on a combination of scores for each event

[a] Education-present cells include text from the corresponding cell above as well. For example, general indication x education-present also includes "Trinase treats and helps prevent seasonal nasal allergy symptoms."

effectiveness is based on a composite score." The item was coded as correct or incorrect based on what information had been present in the stimuli participants had viewed, depending on their assigned experimental condition.

*Composite score comprehension*was measured with four separate items. The first item measured composite score comprehension pertaining to symptoms: "Trinase prevents some but not all seasonal allergy symptoms." Responding "true" (vs. "false") indicated correct comprehension. The second item presented three examples and asked participants to select the one that correctly depicted a composite score. Two additional items measured comprehension of composite score measurement and comprehension of composite score meaning. Respectively, items evaluating these measures asked participants to provide open-ended responses about how scientists measure the effectiveness of Trinase and what "composite score" means as related to prescription drugs. These responses were coded to indicate correct vs. incorrect comprehension.

*Perceived efficacy*was measured with six items. The first item assessed likelihood of benefit ("In your opinion, if 100 people take Trinase, for how many will the drug work?"), where participants chose among six options: 0 people, 20 people, 40 people, 60 people, 80 people, and 100 people. Responses were recorded so that 1 = 0 people through 6 = 100 people. The second item assessed the magnitude of benefit ("In your opinion, if Trinase did help a person's seasonal allergies, how much would it help?") on a scale from 1 (would help allergies a little) to 6 (would help allergies a lot). Further, confidence in drug benefits was measured by combining four items that asked participants to indicate how confident they were that Trinase relieves (1) all seasonal allergy symptoms, (2) congestion, (3) runny nose, and (4) nasal itching on a 5-point scale from 1 (not at all confident) to 5 (extremely confident). The items formed a reliable scale (Cronbach's alpha = 0.86).

*Perceived risk*was measured with two items. One assessed likelihood of risk using a visual scale: "In your opinion, if 100 people take Trinase, what percentage of them will have side effects?" Participants used a slide ruler to choose a percentage between 0 and 100. The second item assessed the magnitude of risk on a 1 (not at all serious) to 6 (very serious) scale: "In your opinion, if Trinase did cause side effects, how serious would they be?"

2.6. Analysis

Analyses were performed on weighted data using SUDAAN 11.0 and the survey procedures in SAS 9.3. Statistical weighting procedures were used to adjust for sampling and survey errors including nonresponse, noncoverage, and under representation of minorities.

For dichotomous outcomes we used logistic regression, running four models for each outcome. Model 1 included type of information as a predictor with composite score definition as a reference category and the educational intervention as a predictor using the education-absent condition as the reference category. Model 2 added the interaction of these two variables. Models 3 and 4 matched the first two models but used the list-of-symptoms condition as the reference category. Significance was defined as $p < 0.05$. We used a Holm-modified Bonferroni adjustment to control for experiment-wise error rates across groups of interaction contrasts [23]. If interaction coefficients were not significant using this adjustment, we focused on results from models without the interactions. For significant interactions, interpretations and application of the Bonferroni–Holm adjustment to simple main effect contrasts rely on methods from Jaccard [24,25]. For ordinal outcome items, we ran a series of ANOVA's with type of information, educational intervention, and an interaction term as the independent variables. We conducted planned comparisons when necessary and applied Bonferroni-adjusted significance levels in these analyses (i.e., $p < 0.017$ [0.05/3 comparisons] for main effects of type of information, and $p < 0.003$ [0.05/15 comparisons] for interactions of type of information and educational intervention).

Table 3
Distributions of outcome variables ($n = 1967$).

Outcome	n (%) or mean (SD)
Correct composite score recognition[a]	1584 (80.5%)
Correct composite score comprehension	
Symptoms	1334 (68.8%)
Example	1763 (92.0%)
Measurement[a]	676 (34.4%)
Meaning[a]	593 (30.2%)
Perceived efficacy	
Likelihood of benefit (1–6 scale)	4.25 (0.91)
Magnitude of benefit (1–6 scale)	4.56 (0.95)
Confidence in drug benefits (1–5 scale)	3.10 (0.87)
Perceived risk	
Likelihood of risk (0–100 scale)	39.02 (23.53)
Magnitude of risk (1–6 scale)	3.91 (1.28)

Note: data are unweighted. Percentages are based on non-missing data.
[a] Participants who did not answer these items were considered to have an incorrect response.

3. Results

3.1. Composite score recognition

Table 3 presents distributions of the outcome variables. Results showed that for type of information, the odds of correctly recognizing that the ad stated that the advertised drug's effectiveness was based on a composite score were higher for those in the composite-score-definition condition than in the list-of-symptoms condition (list of symptoms vs. composite score definition: OR = 0.69, 95% CI: 0.49, 0.99, $p = 0.04$). However, the odds of correctly recognizing that the ad stated that the advertised drug's effectiveness was based on a composite score did not significantly differ between the composite-score-definition condition and the general indication condition. The educational intervention did not affect composite score recognition. The interaction of educational intervention and type of information on composite score recognition was not significant.

3.2. Composite score comprehension

For type of information, the odds of exhibiting composite score comprehension pertaining to symptoms were higher for those in the composite-score-definition condition than in the general indication condition (general indication vs. composite score definition: OR = 0.67, 95% CI: 0.50, 0.91, $p = 0.01$). However, there were no differences in comprehension of symptoms between the composite score definition and the list-of-symptoms conditions. Also, type of information did not affect comprehension of the composite score example. The educational intervention did not significantly affect composite score comprehension pertaining to symptoms or comprehension of the composite score example.

We observed effects of type of information on comprehension of composite score measurement (general indication vs. composite score definition: OR = 0.56, 95% CI: 0.41, 0.77, $p < 0.001$; list of symptoms vs. composite score definition: OR = 0.54, 95% CI: 0.39, 0.74,$p < 0.001$). However, these effects were conditional on educational intervention. Additionally, the educational intervention had an effect on composite score measurement (education-present vs. education-absent: OR = 2.62, 95% CI: 2.02, 3.39, $p < 0.001$), but this effect was contingent on type of information. We observed no main effects of type of information on comprehension of composite score meaning, although the educational intervention did impact comprehension of composite score meaning (education-present vs. education-absent: OR = 1.97, 95% CI: 1.53, 2.54, $p < 0.001$). Again, this effect was contingent on type of information.

The interactions of educational intervention and type of information significantly predicted two comprehension outcomes: composite score measurement and comprehension of composite score meaning. Specifically, the interactions of educational intervention × general indication and educational intervention × list of symptoms were significant in the model predicting comprehension of composite score measurement. The odds ratios for these interactions were 3.42 (95% CI: 1.79, 6.53, $p < 0.001$) and 3.03 (95% CI: 1.58, 5.83, $p = 0.001$), respectively. Table 4 presents the simple main effects of type of information and educational intervention in the model predicting comprehension of composite score measurement. As shown in Block 1 of the table, among those exposed to the list-of-symptoms or the general indication conditions, the odds of comprehending composite score measurement were higher if the educational intervention was present than if it was absent. However, the educational intervention did not affect this comprehension outcome among those assigned to the composite-score-definition condition. Block 2 shows that without the educational intervention, seeing the composite score definition yielded higher odds of comprehension than did seeing the general indication or list of symptoms.

In the model predicting comprehension of composite score meaning, the interaction of educational intervention × list of symptoms was significant (OR = 2.24, 95% CI: 1.20, 4.16, $p = 0.01$). Table 5 presents the simple main effects of type of information and educational intervention in the model predicting comprehension of composite score meaning. Results are similar to those seen for comprehension of composite score measurement. The educational intervention mattered for comprehending composite score meaning when the composite score definition was not available (Block 1), and the composite score definition mattered for this outcome if the educational intervention was not available (Block 2).

3.3. Perceived efficacy

The type of information and educational intervention did not have an effect on perceived efficacy likelihood or perceived efficacy magnitude. However, there were main effects of type of information ($F(2, 1951) = 30.92, p < 0.001$) and educational intervention ($F(1, 1951) = 6.03, p = 0.01$) on confidence in drug benefits. Participants who saw the composite score definition ($M = 2.91$, SE = 0.04) reported significantly less confidence in the drug's benefits than participants in the general indication ($M = 3.12$, SE = 0.04) and list-of-symptoms conditions ($M = 3.39$, SE = 0.05). In addition, participants who saw the educational intervention ($M = 3.08$, SE = 0.04) had less confidence in the drug's benefits than participants who did not see the educational intervention ($M = 3.20$, SE = 0.04).

None of the interactions were significant.

3.4. Perceived risk

The type of information and the educational intervention did not have an effect on perceived risk likelihood. The type of

Table 4
Simple main effect contrasts of education intervention and type of information on comprehension of composite score measurement.

Block	Comparison groups	Odds ratio	95% CI		p Value
1	Education present vs. education absent for the composite score definition group	1.28	0.85	1.92	0.23
	Education present vs. education absent for the list-of-symptoms group	3.88	2.33	6.48	<.001[a]
	Education present vs. education absent for the general indication group	4.38	2.65	7.24	<.001[a]
2	General indication vs. composite score definition for the education-absent group	0.28	0.17	0.47	<.001[a]
	List of symptoms vs. composite score definition for the education-absent group	0.29	0.17	0.49	<.001[a]
	General indication vs. list of symptoms for the education-absent group	0.97	0.53	1.78	0.92
3	General indication vs. composite score definition for the education-present group	0.97	0.65	1.44	0.86
	List of symptoms vs. composite score definition for the education-present group	0.88	0.59	1.33	0.55
	General indication vs. list of symptoms for the education-present group	1.09	0.75	1.60	0.65

[a] Statistically significant using the Holm-modified Bonferroni adjustment. Each block represents a separate family of contrasts to which we applied this adjustment.

Table 5
Simple main effect contrasts of education intervention and type of information on comprehension of composite score meaning.

Block	Comparison groups	Odds ratio	95% CI		*p* Value
1	Education present vs. education absent for the composite score definition group	1.32	0.86	2.02	0.20
	Education present vs. education absent for the list-of-symptoms group	2.95	1.88	4.64	<.001[a]
	Education present vs. education absent for the general indication group	2.05	1.30	3.24	0.002[a]
2	General indication vs. composite score definition for the education absent group	0.68	0.42	1.10	0.11
	List of symptoms vs. composite score definition for the education absent group	0.56	0.35	0.88	0.01[a]
	General indication vs. list of symptoms for the education absent group	1.23	0.74	2.04	0.43
3	General indication vs. composite score definition for the education present group	1.06	0.71	1.59	0.78
	List of symptoms vs. composite score definition for the education present group	1.24	0.82	1.88	0.30
	General indication vs. list of symptoms for the education present group	0.85	0.58	1.26	0.43

[a] Statistically significant using the Holm-modified Bonferroni adjustment. Each block represents a separate family of contrasts to which we applied this adjustment.

information did have a main effect on perceived risk magnitude ($F(2, 1952) = 3.60$, $p = 0.03$). Post-hoc comparisons indicated that perceived risk magnitude scores were significantly lower in the composite-score-definition condition ($M = 3.76$, SE = 0.06) than in the general indication condition ($M = 3.99$, SE = 0.06); there were no significant differences in this outcome between the list-of-symptoms condition and the other types of information conditions.

Although there was no main effect for the educational intervention, the type of information and educational intervention had a significant interactive effect on perceived risk magnitude ($F(2, 1952) = 3.23$, $p = 0.04$). The significant interaction appeared to be driven by two sets of comparisons: First, perceived risk magnitude was significantly lower in the composite score definition education-present condition ($M = 3.71$, SE = 0.09) than in the general indication education-absent condition ($M = 4.18$, SE = 0.10). Second, perceived risk magnitude was significantly lower in the general indication education-present condition ($M = 3.81$, SE = 0.07) than in the general indication education-absent condition ($M = 4.18$, SE = 0.10).

4. Discussion and conclusion

4.1. Discussion

To test which combinations maximize consumer comprehension and informed decision-making, this experimental study examined how individuals interpret and respond to mock DTC prescription drug ads that varied their presentation of composite endpoint efficacy information. Ads varied by the type of information featured (general indication information, a list of symptoms, or a definition of composite scores) and whether the information was accompanied by an educational intervention to teach consumers about composite scores. The overall pattern of findings from the study suggests that either presenting information accompanied by a composite score definition or providing an educational intervention improves composite score comprehension, but there was no additional value of including both.

The educational intervention improved composite score comprehension *only* when it was paired with the list-of-symptoms or the general indication condition. However, when paired with the composite-score-definition condition, the educational intervention did not improve composite score comprehension. This finding is not surprising because the ad featuring the combination of the composite score definition with the educational intervention contained a lot of information to read and process and may have required too much cognitive effort [26]. The overall pattern of findings suggested that the composite-score-definition condition was the most effective type of information at improving both recognition that the ad stated that the advertised drug's effectiveness was based on a composite score and composite score comprehension.

Accurate perceptions of risk and efficacy are central to making informed decisions about prescription drugs. Yet, some research has found that DTC ads can lead consumers to overestimate drug benefits and underestimate drug risks [7–10]. We found that participants in the composite-score-definition condition had lower perceptions of the magnitude of the drug risk than participants in the general indication condition but not the list-of-symptoms condition. However, participants in the composite-score-definition condition also had less confidence in the drug's benefits than participants in the general indication and list-of-symptoms conditions. Similarly, participants who saw the educational intervention had less confidence in the drug's benefits, and participants in the composite-score-definition and general indication conditions who saw the educational intervention had lower perceptions of the magnitude of the drug risk, than participants who did not see the educational intervention. This finding suggests that the composite score efficacy information in these ads may have helped participants overcome drug efficacy, but not drug risk, misperceptions frequently found in response to DTC ads [7–10]. The finding is not surprising given that the composite score efficacy information in these ads focuses on drug efficacy, not drug risk. Further there was direct correspondence between the information conveyed (particularly in the composite-score-definition condition) and the way that confidence in the drug's benefits was measured. Specifically, the composite-score-definition condition was the only condition that informed participants that the drug "may not be as effective in addressing each symptom individually," and confidence was measured by asking participants how confident they were about the drug's ability to relieve (1) all seasonal allergy symptoms, (2) congestion, (3) runny nose, and (4) nasal itching. Because participants had just read that the drug may not be effective for individual symptoms, it only makes sense that they would respond that they had less confidence in the three items that measured individual symptoms.

This study had some methodological limitations. First, we presented the ads online to members of an Internet panel. As such, although they could view it as long as they wished once, participants only saw the ad once, rather than having the opportunity to return to the ad for additional viewing as they could in the real world. Also, we examined only print ads that display a static message because this was a foundational study, the first of its kind exploring composite score comprehension. However, because DTC ads are displayed through a range of popular media such as magazines, newspapers, Internet, social media, and television, future studies should investigate media differences to determine whether the findings generalize to these different media.

In addition, accurately measuring the complex concept of composite score comprehension was challenging. We measured comprehension using two closed-ended questions and two open-ended items. The majority of participants provided correct responses to the closed-ended items; in particular, the closed-ended item that asked participants to identify the example of a composite score from a list had little variability. In contrast, the majority of participants provided responses to the open-ended items that indicated incorrect comprehension. It appeared to be especially difficult for participants to define the term "composite score" (e.g., responding "don't know" or providing irrelevant responses). Because the open-ended items were hard for participants, their responses yielded little information that could be used to develop additional response options for future closed-ended items. Further development and testing is needed to identify the best way to measure composite score comprehension.

4.2. Conclusion

Drug efficacy is often measured by multiple endpoints that are then combined into an overall composite score [15,16]. Yet, DTC ads rarely highlight the fact that their benefits are the combination of endpoints that may not all show promising effects. This is the first study to examine whether consumers notice composite scores in DTC ads and whether they understand how composite scores are used for measuring drug efficacy. We tested three different ways of presenting the information that arises from a composite score and an educational intervention to explain the concept of a composite score. Of the six combined presentation styles that we examined, two options are worthy of further consideration. One option is to present the definition of composite scores (composite-score-definition condition) alone. The other option is to include the drug indication or the list of symptoms treated along with an explanation of composite scores (i.e., list-of-symptoms or general indication condition along with an educational intervention using a description similar to the decathlon example, as seen in Table 2). These recommendations are consistent with the literature that suggests limiting the amount of information presented in order to facilitate comprehension [27,28]; these presentations of composite score information provide succinct yet sufficient explanations of the concept. Given that there were no robust findings suggesting that any one combination of information was superior to the others, it would be important to test these recommendations further before implementing policy changes based on these findings.

4.3. Practice implications

Effectively delivering composite score information may help patients understand drug efficacy information based on composite scores, which may ultimately lead to patients making more informed decisions. We are unaware of any DTC ads that acknowledge when advertised benefits result from a combination of endpoints, perhaps as a result of consumers' general lack of familiarity with this concept. In order to explore whether this information would benefit consumers, appropriate measurement issues need to be addressed. Future research should focus on developing an accurate way to measure composite score comprehension. Additional work should be done to test the effect of different ways of presenting composite score information on print ads using the two suggested options mentioned above along with other similar conditions, as well as with different modes of media. Future research will need to be directed toward specifying the importance of including an example of a composite score: the primary difference between the two options recommended here is that the list-of-symptoms or general indication condition combined with the educational intervention provides an everyday example of composite scores, while the composite-score-definition condition alone does not. Additional research might also explore insights into the process of understanding composite scores and their relations to drug efficacy and risk. For instance, by conducting cognitive interviews employing "think aloud" techniques and probing questions after participants have viewed ads with different ways of presenting composite score information, researchers would be able to gain insight into how patients figure out the connection between composite score information and the advertised drug's efficacy and risk.

Acknowledgements

The Center for Drug Evaluation and Research, Food and Drug Administration (FDA) provided research funding for this study. The authors thank Chelsea Burfeind of RTI International for research assistance and Kayla Gray and Maria Ashbaugh of RTI International for their assistance with stimuli development.

References

[1] IMS Institute for Healthcare Informatics, The Use of Medicines in the United States: Review of 2011, IMS Institute for Healthcare Informatics, Parsippany, NJ, 2012 Available from: http://www.imshealth.com/ims/Global/Content/Insights/IMS%20Institute%20for%20Healthcare%20Informatics/IHII_Medicines_in_U.S_Report_2011.pdf (accessed August, 2014).
[2] Centers for Medicare & Medicaid Services, National Health Expenditure Data Fact Sheet, Centers for Medicare & Medicaid Services, Baltimore, MD, 2015 Available from: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.html, (updated 03.12.14, cited 05.11.14).
[3] R.A. Bell, R.L. Kravitz, M.S. Wilkes, Direct-to-consumer prescription drug advertising, 1989–1998. A content analysis of conditions, targets, inducements, and appeals, J. Fam. Pract. 49 (2000) 329–335 PubMed PMID: 10778839.
[4] S. An, Attitude toward direct-to-consumer advertising and drug inquiry intention: the moderating role of perceived knowledge, J. Health Commun. 12 (2007) 567–580 PubMed PMID: 17763053.
[5] J.J. Davis, Riskier than we think? The relationship between risk statement completeness and perceptions of direct to consumer advertised prescription drugs, J. Health Commun. 5 (2000) 349–369.
[6] L.H. Glinert, J.C. Schommer, Television advertisement format and the provision of risk information about prescription drug products, Res. Soc. Adm. Pharm. 1 (2005) 185–210.
[7] J.J. Davis, Consumers' preferences for the communication of risk information in drug advertising, Health Aff. (Millwood) 26 (2007) 863–870 PubMed PMID: 17485767.
[8] D.L. Frosch, P.M. Krueger, R.C. Hornik, P.F. Cronholm, F.K. Barg, Creating demand for prescription drugs: a content analysis of television direct-to-consumer advertising, Ann. Fam. Med. 5 (2007) 6–13 PubMed PMID: 17261859.
[9] K.A. Kaphingst, W. DeJong, R.E. Rudd, L.H. Daltroy, A content analysis of direct-to-consumer television prescription drug advertisements, J. Health Commun. 9 (2004) 515–528 PubMed PMID: 15764450.
[10] J. Kees, P.F. Bone, J. Kozup, P.S. Ellen, Barely or fairly balancing drug risks? Content and format effects in direct-to-consumer online prescription drug promotions, Psychol. Market. 25 (2008) 675–691 PubMed PMID: WOS: 000258275400008.
[11] A. Lyles, Direct marketing of pharmaceuticals to consumers, Annu. Rev. Public Health 23 (2002) 73–91 PubMed PMID: 11910055.
[12] J.P. Griffin, B.M. Godfrey, R.E. Sherman, Regulatory requirements of the food and drug administration would preclude product claims based on observational research, Health Aff. (Millwood) 31 (2012) 2188–2192 PubMed PMID: 23048095.
[13] A.S. Kesselheim, J. Avorn, The food and drug administration has the legal basis to restrict promotion of flawed comparative effectiveness research, Health Aff. (Millwood) 31 (2012) 2200–2205 PubMed PMID: 23048097.
[14] G.H. Rutan, R.H. McDonald, L.H. Kuller, A historical perspective of elevated systolic vs diastolic blood pressure from an epidemiological and clinical trial viewpoint, J. Clin. Epidemiol. 42 (1989) 663–673 PubMed PMID: 2668449.
[15] Agency for Healthcare Research and Quality (AHRQ), Combining Measures into Composite or Summary Scores, Agency for Healthcare Research and Quality, Rockville, MD, 2012 Available from: https://www.talkingquality.ahrq.gov/content/create/scores/combinemeasures.aspx, (accessed 05.11.14).
[16] American Medical Association (AMA), Measures Development, Methodology, and Oversight Advisory Committee: Recommendations to PCPI Work Groups on Composite Measures, American Medical Association, Chicago, IL, 2010

Available from: http://www.ama-assn.org/resources/doc/cqi/composite-measures-framework.pdf, (accessed 05.11.14).
[17] A.C. O'Donoghue, H.W. Sullivan, P.A. Williams, C.M. Squire, K.R. Betts, J.F. Willoughby, S. Parvanta, (under review). Consumers' Understanding of FDA Approval Requirements and Composite Scores in Direct-to-Consumer Prescription Drug Print Ads.
[18] L. Chang, J.A. Krosnick, National surveys via RDD telephone interviews versus the Internet: comparing sample representativeness and response quality, Public Opin. Q. 73 (2009) 641–678.
[19] M. Dennis, Knowledge Panel: Processes and Procedures Contributing to Sample Representativeness and Tests for Self-selection Bias, Knowledge Networks, Menlo Park, CA, 2010, pp. 2010 Available from: http://www.knowledgenetworks.com/ganp/docs/KnowledgePanelR-Statistical-Method.
[20] A.C. O'Donoghue, H.W. Sullivan, K.J. Aikin, D. Chowdhury, R.R. Moultrie, D.J. Rupert, Presenting efficacy information in direct-to-consumer prescription drug advertisements, Patient Educ. Couns. 95 (2014) 271–280 PubMed PMID: 24581929.
[21] A.C. O'Donoghue, H.W. Sullivan, K.J. Aikin, Randomized study of placebo and framing information in direct-to-consumer print advertisements for prescription drugs, Ann. Behav. Med. 48 (2014) 311–322 PubMed PMID: 24682975.
[22] A.C. O'Donoghue, P.A. Williams, H.W. Sullivan, V. Boudewyns, C. Squire, J.F. Willoughby, Effects of comparative claims in prescription drug direct-to-consumer advertising on consumer perceptions and recall, Soc. Sci. Med. 120C (2014) 1–11 PubMed PMID: 25194471.
[23] J. Jaccard, R. Turrisi, Interaction effects in multiple regression, Sage University Papers Series: Quantitative Applications in the Social Sciences (07-072), 2nd ed., Sage University, Thousand Oaks, CA, 2003.
[24] J. Jaccard, Interaction effects in factorial analysis of variance, Papers Series: Quantitative Applications in the Social Sciences (07-118), Sage University, Thousand Oaks, CA, 1998.
[25] J. Jaccard, Interaction effects in logistic regression, Sage University Papers Series: Quantitative Applications in the Social Sciences (07-135), Sage University, Thousand Oaks, CA, 2001.
[26] A.D. Gurmankin, J. Baron, K. Armstrong, The effects of numerical statements of risk on trust and comfort with hypothetical physician risk communication, Med. Decis. Making 24 (2004) 265–271.
[27] A. Fagerlin, E. Peters, Quantitative information, in: B. Fishoff, N. Brewer, J. Downs (Eds.), Communicating Risks and Benefits: An Evidence-based User Guide, Food and Drug Administration, U.S. Department of Health and Human Services, Washington, D.C, 2011.
[28] E. Peters, D. Vastfjall, P. Slovic, C.K. Mertz, K. Mazzocco, S. Dickert, Numeracy and decision making, Psychol. Sci. 17 (2006) 407–413 PubMed PMID: 16683928.