Taylor & Francis
Taylor & Francis Group

Journal of Health Communication
International Perspectives

ISSN: 1081-0730 (Print) 1087-0415 (Online) Journal homepage: http://www.tandfonline.com/loi/uhcm20

Comprehensibility of Health-Related Documents for Older Adults with Different Levels of Health Literacy: A Systematic Review

Ruth Koops van 't Jagt, John C. J. Hoeks, Carel J. M. Jansen, Andrea F. de Winter & Sijmen A. Reijneveld

To cite this article: Ruth Koops van 't Jagt, John C. J. Hoeks, Carel J. M. Jansen, Andrea F. de Winter & Sijmen A. Reijneveld (2016) Comprehensibility of Health-Related Documents for Older Adults with Different Levels of Health Literacy: A Systematic Review, Journal of Health Communication, 21:2, 159-177

To link to this article: http://dx.doi.org/10.1080/10810730.2015.1049306

Journal of Health Communication, 21:159–177, 2016

ISSN: 1081-0730 print/1087-0415 online
DOI: 10.1080/10810730.2015.1049306

Routledge
Taylor & Francis Group

Comprehensibility of Health-Related Documents for Older Adults with Different Levels of Health Literacy: A Systematic Review

RUTH KOOPS VAN 'T JAGT[1], JOHN C. J. HOEKS[1], CAREL J. M. JANSEN[1], ANDREA F. DE WINTER[2], and SIJMEN A. REIJNEVELD[2]

[1]*Department of Communication and Information Sciences, Faculty of Arts, University of Groningen, Groningen, The Netherlands*
[2]*Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands*

A systematic review was conducted to assess the available evidence for the effectiveness of interventions aiming to improve the comprehensibility of health-related documents in older adults (≥50) with different levels of health literacy. Seven databases were searched (2005 forward), and references in relevant reviews were checked. The selection procedure was conducted by 2 independent reviewers. Data extraction and assessment of the quality of the resulting studies were conducted by 1 reviewer and checked for accuracy by a 2nd reviewer. A total of 38 intervention studies had a study population of older adults ($n = 35$) or made an explicit comparison between age groups, including older adults ($n = 3$). Inconsistent evidence was found for the importance of design features to enhance the comprehensibility of health-related documents. Only for narratives and multiple-feature revisions (e.g., combining revisions in textual and visual characteristics) did the included studies provide evidence that they may be effective for older adults. Using narrative formats and/or multiple-feature revisions of health-related documents seem to be promising strategies for enhancing the comprehensibility of health-related documents for older adults. The lack of consistent evidence for effective interventions stresses the importance of (a) replication and (b) the use of standardized research methodologies.

Older adults (i.e., those aged 50 and older) are frequently affected by the negative consequences of limited health literacy (Zamora & Clingerman, 2011). *Health literacy* can be defined as the degree to which people are able to access, understand, appraise, and communicate information in order to engage with the demands of different health contexts so as to promote and maintain health across the life course (Kwan et al., 2006). Low health literacy is found in 36% to 68% of older adults (Adams et al., 2013; Ashida et al., 2011; Jovic-Vranes & Bjegovic-Mikanovic, 2012; Ownby, Waldrop-Valverde, & Taha, 2012). Low levels of health literacy have frequently been associated with poor health outcomes (Al Sayah, Majumdar, Williams, Robertson, & Johnson, 2013; Kim, 2009; Mõttus et al., 2014; Omachi, Sarkar, Yelin, Blanc, & Katz, 2013).

Older adults with limited health literacy have difficulty understanding health documents such as instructions, medication labels, patient education materials, consent forms, and health surveys. The appropriate use and comprehension of these health-related documents in prevention, care, and cure settings is crucial for older adults for access to and utilization of health care and management of health and illness (Bostock & Steptoe, 2012; Morrow et al., 2006; Wolf, Gazmarian, & Baker, 2005). Therefore, the focus of this review is on the comprehensibility of health-related documents. Health care professionals and policymakers are increasingly aware of the importance of appropriate health-related documents but lack knowledge of the formats and features of these documents that can contribute to the comprehensibility of health information (Coleman, Hudson, & Maine, 2013).

Reviews of evidence for document design interventions aimed at enhancing comprehensibility (e.g., Berkman et al., 2011; Sheridan et al., 2011) have concluded that the strength of evidence and generalizability of findings are often low, largely because of heterogeneity regarding the type of interventions and type of populations studied. Furthermore, as these reviews have not focused specifically on older adults, researchers cannot draw firm conclusions about formats and features that may enhance comprehensibility for this group. Aging is associated with deteriorations in cognitive abilities that are essential for the comprehension of health-related and other documents. In older adults these cognitive abilities may be negatively affected by limitations in processing speed and working memory (cf. Chin et al., 2011). So far it has not been determined which characteristics of health-related documents can and should be influenced that help or hinder in enhancing comprehensibility

Address correspondence to Ruth Koops van 't Jagt, Department of Communication and Information Sciences, Faculty of Arts, University of Groningen, P.O. Box 716, Groningen 9700 AS, The Netherlands. E-mail: r.koops.van.t.jagt@rug.nl

in older adults, especially in those with limited health literacy. Insight into these characteristics is essential to help health care professionals and policymakers improve health communication. The objective of this study is to systematically review the evidence for the effectiveness of interventions that aim to improve the comprehensibility of health-related documents in older adults, with special attention to the effect of health literacy.

Methods

Search Strategy

We conducted a systematic search of original research studies, systematic reviews, and nonsystematic reviews of interventions that aim to improve the comprehensibility of health-related documents. We focused on studies that at least included a subgroup of older adults, defined as persons 50 years of age and older. There is evidence that age-related differences in health literacy exist between middle-aged adults (e.g., those 45–59 years of age) and younger age groups (HLS-EU Consortium, 2012; Jovic-Vranes & Bjegovic-Mikanovic, 2012). We used MEDLINE, PsycINFO, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science (WoS), The Cochrane Library, Educational Resources Information Center (ERIC), and the Comprehensible Language and Effective Communication (CLEC) database. Reference lists of key articles were manually searched to identify further relevant articles. All databases were searched for publications dating from January 1, 2005, to March 7, 2014. Search terms consisted of terms related to health literacy, to health-related documents, and to comprehensibility (see Appendix A for search strategy results in MEDLINE).

Selection of Studies

The selection of studies was conducted in three separate phases: title review, abstract review, and full-text review. In the title review phase, all references were screened by one researcher (Reviewer A). A high-tolerance strategy was applied, excluding only those titles that were clearly not relevant (such as "Speaking Up: Teens Voice Their Health Information Needs"; Smart, Parker, Lampert, & Sulo, 2012). In the next phase, the abstracts of articles resulting from the title review were screened for relevance by pairs of reviewers (Reviewer A and B, C, or D). Finally, the full texts of articles were screened for relevance during the full-text review phase by one reviewer and checked for accuracy by a second reviewer (Reviewers A and B or Reviewers A and C). Articles were included if they provided information on the effectiveness of interventions aiming to improve the comprehensibility of health-related documents in older adults from industrialized countries, with special attention to possible effects of health literacy. The inclusion and exclusion criteria are shown in Table 1.

Full-Text Review: Data Extraction and Quality Assessment

A coding form was developed for data extraction. This form captured general information, the main research question, methodological data, characteristics of the included populations, data about the health-related documents and the interventions reported on, the results and conclusions as reported by the authors. Data were assessed by one reviewer and checked for accuracy by a second reviewer (Reviewers A and B or Reviewers A and C). Discrepancies were resolved via discussion and consultation with team members.

The studies identified were classified in accordance with the communication topics defined by Abraham and Kools (2012):

1. Studies that focus on the effectiveness of different *media formats* of health-related documents
2. Studies that address the design of *presentation of information* within one medium
   a. Graphical formats (e.g., presentation of numerical information)
   b. Pictures
   c. Textual design (e.g., order and layout)

Table 1. Inclusion and exclusion criteria

Data type	Primary research, quantitative data
Participants	Studies including participants 18 years of age and older. Studies were excluded if (a) they focused solely on children, adolescents, or young adults; (b) the results for older adults (ages $\geq$50) were not provided separately from the results for adults of other age groups; (c) they focused on a nonrelevant subgroup (e.g., parents, military personnel).
Setting	All studies that took place in Western countries were included.
Intervention	Any single or complex intervention in which at least one feature of health-related documents was manipulated or varied. Studies were excluded if (a) documents aimed to measure the health-related knowledge of people, (b) documents concerned informed consent forms for research purposes, or (c) the study focused on relevant health documents aimed at health professionals who work with older adults (because the focus of this review is on older adults themselves).
Outcomes	Any outcomes and measures that were regarded as acceptable as an indicator of either comprehension or comprehensibility (a full list of outcomes and measures regarded as acceptable is provided in Appendix B).
Study design	All study designs and assessments among participants. Studies were excluded if they only applied comprehensibility analysis with a readability formula.

Table 2. Checklist used to assess the quality of studies (modified from Downs & Black, 1998)

Quality item	Criteria
Reporting	Eight items 1. Hypothesis/aim/objectives clearly described 2. Outcomes to be measured described in Introduction or Method section 3. Characteristics of participants clearly described 4. Interventions clearly described 5. Distribution of principal confounders clearly described 6. Main findings clearly described 7. Estimates of random variability provided 8. Characteristics of excluded participants provided
External validity (to address the representativeness of the findings of the study for the population the study focused on, for health-related documents in general, and for situations in which older people may encounter health-related documents)	Three items 9. Study participants representative 10. Health-related documents representative 11. Study setting representative (e.g., no lab setting)
Internal validity—investigator bias	Four items 12. Statistical tests appropriate 13. Outliers reported 14. Blinding of researchers appropriate 15. All analyses planned ahead
Internal validity—selection bias	Three items 16. Participants from the same population distributed over different comparison groups 17. Randomization to study groups undertaken 18. Adjustment for confounding undertaken if necessary
Internal validity—construct validity and measurement validity/reliability (to assess whether the constructs and measurement instruments used in the study represented the underlying concepts in a scientifically acceptable manner)	Three items 19. Measurements reliable 20. Outcome measures valid and reliable 21. Manipulations valid

3. Studies that address *linguistic characteristics* of health-related documents
4. Studies that address the effectiveness of *multiple-feature revisions* (including revisions of various features of design of presentation of information and/or linguistic characteristics)
5. Studies that address the *design* of informational content in health-related documents
   a. Message framing
   b. Narratives (stories)
6. Studies that address *other factors* that may contribute to comprehensibility, such as presence of an external aid, or learning method applied.

The quality of the studies was assessed using a checklist based on Downs and Black (1998; see Table 2). The final checklist consisted of 21 items. Quality scores between 18 and 21 were considered high (HQ), scores between 15 and 17 fair (FQ), scores between 12 and 14 marginal (MQ), and scores less than 12 poor (PQ). Quality was assessed by one reviewer and checked for accuracy by a second reviewer.

Analysis and Reporting

Quantitative analysis or statistical pooling of the data was considered inappropriate because of the wide variety of outcome measures used and the diversity of interventions reported on. Therefore, a narrative data synthesis was conducted based on the framework proposed by the Economic and Social Research Council Guidance Project (Popay et al., 2006).

First, we report descriptive data for the studies that were included. After assessing their quality, we summarize findings regarding effects by type of intervention (classified in accordance with the communication topics defined by Abraham & Kools, 2012). Finally, we summarize evidence for the effects for subgroups that differ in their levels of health literacy.

Table 3 shows the way in which we assessed levels of evidence. Because no standard approach exists for assessing levels of evidence in document intervention studies, we defined levels of evidence based on a classification for other intervention studies (e.g., Creemers, Verhulst, & Huizink, 2009) that we adapted for the purpose of the current review. Thus, an adapted best evidence synthesis, as proposed by Van Tulder, Furlan, Bombardier, and Bouter (2003), was performed. The quality of studies was taken into account by only including studies of at least fair quality. No additional weighting was done of studies of fair or high quality, and levels of evidence did not depend on whether the evidence for the intervention benefit was positive or negative.

Table 3. Levels of evidence

Level of evidence	Criteria
Strong evidence	A large majority (≥75%) of four or more fair- to high-quality studies report consistent findings for the effectiveness of a specific feature or form of health-related documents.
Moderate evidence	A majority (≥65%) of two to three fair- to high-quality studies report consistent findings for the effectiveness of a specific feature or form of health-related documents.
Weak evidence	A small majority (≥60%) of four or more fair- to high-quality studies report consistent findings for the effectiveness of a specific feature or form of health-related documents.
Inconsistent evidence	Inconsistent findings in studies of fair to high quality: Some studies report evidence for the effectiveness of a specific feature or form of health-related documents, whereas other fair- to high-quality studies report no differences in the effectiveness of health-related documents varying in formats or features.
Inconclusive evidence	No studies of fair to high quality or only one study of fair to high quality is available that reports findings for the effectiveness of a specific feature or form of health-related documents.

Results

Search Results

Figure 1 presents a complete overview of the selection process. Four reviewers were involved in the abstract inclusion phase, with each abstract reviewed by two reviewers. Overall, inter-reviewer agreement on abstract inclusion was high (90%–91%), and Cohen's kappas were good (0.61–0.74). Disagreements, mostly resulting from missed details, were resolved by discussion. A total of 38 studies were included that focused on older adults ($n = 35$) or made an explicit comparison between a younger and an older

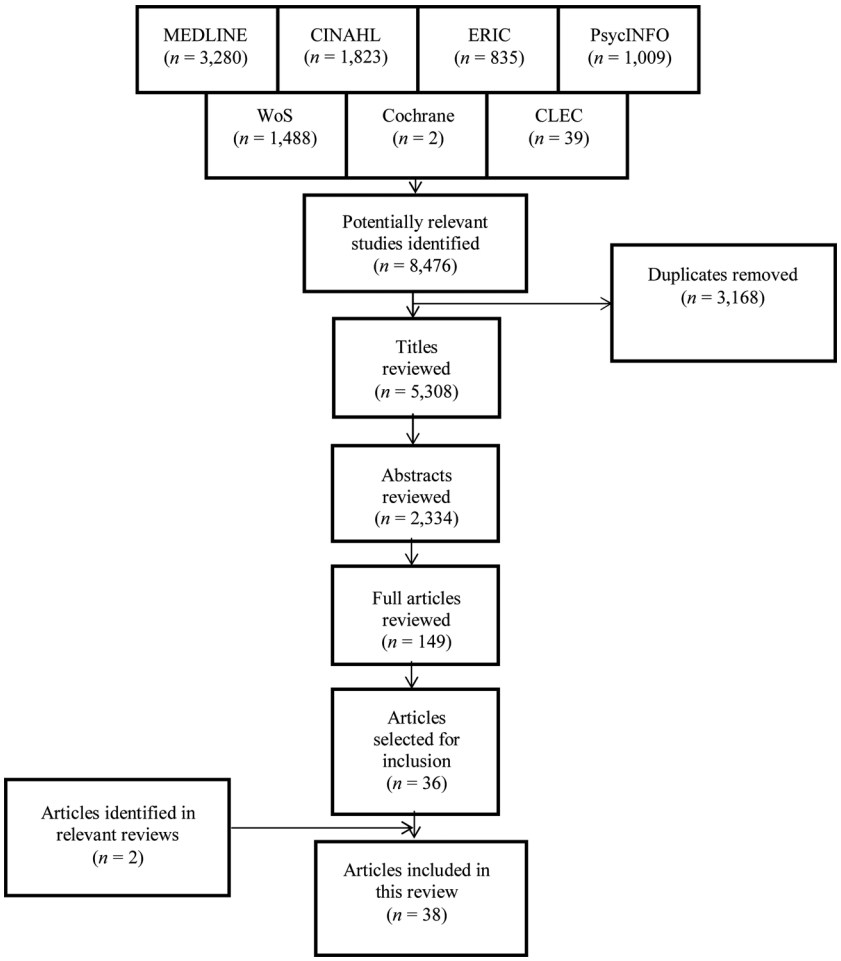

Figure 1. Flowchart of the selection procedure. CINAHL = Cumulative Index to Nursing and Allied Health Literature; ERIC = Educational Resources Information Center; WoS = Web of Science; CLEC = Comprehensible Language and Effective Communication.

subgroup ($n = 3$). Here we provide an overview of these 38 studies. To ensure that we did not miss any important papers, we performed a quick scan of the 808 references from earlier than 2005 in our included articles, which resulted in 15 additional relevant papers, which were then screened for results and conclusions. These additional papers did not alter the conclusions from the articles included in this review.

Characteristics of the Included Studies

A wide variety of topics was addressed, including information about different diseases (cancer, arthritis, anticoagulation, stroke, diabetes, etc.), medication instructions, package inserts, information about care providers (hospitals, home care providers, health care plans, etc.), and information about the evaluation of online health information. The main purposes of the documents were to inform, to educate, and to help people decide (e.g., between treatment options). The online appendix (Table A1) provides an overview of the included studies, their characteristics, and their main outcomes. The main findings are summarized here according to the categorization of communication topics as mentioned in ''Full-Text Review: Data Extraction and Quality Assessment''; some studies addressed multiple issues and were classified accordingly under multiple topics. Studies of all quality rates are discussed here. Table 4 shows the overall levels of evidence for all topics. Table 5 shows the primary outcome measures regarding comprehensibility and other (secondary) outcome measures (such as preferences, intentions, and behavior) used in the included studies.

Quality Assessment

An overview of the results of the study quality assessment is presented in Table 6.

Summary of Effects by Intervention Type

Eight studies compared using multimedia formats (addressing various senses by combining visual and audio information) to using single media formats (addressing only one sense; e.g., by using only a sound clip or only a booklet). Only one study found better comprehension for the multimedia format (Frosch, Legare, & Mangione, 2008, FQ). The other studies found mixed results (Shukla, Daly, & Legutko, 2012, FQ) or no differences (Astley, Chew, Aylward, Molloy, & De Pasquale, 2008, FQ; Ilic, Egberts, McKenzie, Risbridger, & Green, 2008, HQ; Mittal et al., 2007, HQ; Volk et al., 2008, HQ; Xie, 2011, HQ). One study found worse comprehension for the multimedia format (Gattellari & Ward, 2005, HQ). Taken together, this may be considered weak evidence for the absence of effects on comprehension between using single and multimedia formats.

Eight studies assessed the possible effect of different ways of presenting information within one medium, that is, using graphical formats, pictures, and different forms of textual design. Three studies assessed the added value

Table 4. Overall level of evidence for the effectiveness of features and formats of health-related documents on comprehension in older adults

Feature and format	No. of fair- to high-quality studies/ no. of total studies	Distribution of direction of effects in studies of fair to high quality				Level of evidence
		POS	NEG	NULL	MIX	
Media formats						
Multimedia formats	8/8	1	1	5	1	Weak
Presentation of information						
Graphical formats	2/3	1	0	0	1	Inconsistent
Pictures	1/4	0	0	1	0	Inconclusive
Order	0/1					Inconclusive
Linguistic characteristics						
Simplifying language	4/5	1	0	1	2	Inconsistent
Other linguistic characteristics	2/3	1	0	1	0	Inconsistent
Multiple-feature revisions						
Multiple-feature revisions	5/6	3	0	1	1	Weak
Informational content						
Framing	1/3	0	0	1	0	Inconclusive
Narratives	2/2	2	0	0	0	Moderate
Other factors						
Learning method	1/1	0	0	1	0	
Type of computer assistant	0/1					Inconclusive
Descriptors of numerical information	1/2	1	0	0	0	Inconclusive
External aid	2/2	1	0	1	0	Inconsistent
Total studies	27/38					

Note. POS, NEG, NULL, MIX = the effectiveness of the feature or format: positive (POS), negative (NEG), no difference between conditions (NULL), inconsistent or mixed findings (MIX). Two studies were classified in two categories.

Table 5. Primary and secondary outcome measures by study, grouped by type of outcome

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
Knowledge			
Ruiz et al. (2013)	Multimedia computer-based tutorial about cardiovascular risk	Risk understanding measured by gist and verbatim knowledge (understanding one's own risk and precise numerical representation of events) Risk understanding measured by raw data question and by frequency and percent change question Measured at Time 1 as risk understanding, measured at Time 2 as short-term recall, measured at Time 3 as long-term recall	Confidence Perception of importance Perception of seriousness Intent to adhere to risk factor modification Self-efficacy in performing risk factor modification Actual adherence to risk factor modification Accessibility of the information Attitudes toward the computer program
Frosch et al. (2008)	Information about prostate or colon cancer	Knowledge about prostate or colon cancer	Role preference Attitudes toward shared decision making Perceived social norms Self-efficacy Cancer screening decisions
Volk et al. (2008)	Decision aid on prostate cancer and screening	Knowledge of prostate cancer and screening (questions with options of a "yes," "no," or "unsure" response; the content of the questions was drawn from the Interactive Learning Modules and from the factual information in the audio booklet)	Acceptability of the decision aids Engagement with the entertainment-based aid Decisional conflict scale Patient self-advocacy scale (patient involvement)
Gattellari & Ward (2005)	Information about prostate cancer	Knowledge (14-item knowledge measure composed of 10 true/false questions and four multiple-choice questions administered at pretest and posttest reflecting what men should know before undergoing PSA screening as identified by Australian expert consensus)	Men's views toward PSA screening Decisional conflict Decisional control Worry about prostate cancer Perceived ability to make an informed choice about PSA screening Propensity to undergo PSA screening Likelihood of accepting a doctor's recommendation to undergo PSA screening Scenario-based assessment of the appropriateness of two different approaches to PSA screening in general practice Men's perceptions of GP fault regarding adverse consequences of screening decisions Evaluation of materials received
Ilic et al. (2008)	Information about prostate cancer	Knowledge (a five-item multiple-choice questionnaire assessing knowledge about prostate cancer and PSA testing; for each participant, the percentage of items answered correctly was calculated)	Decisional conflict Anxiety Consumer decision-making role Screening interest
Zikmund-Fisher et al. (2008)	Information about risks and benefits of medication	Gist knowledge (essential knowledge of the side effects discussed in the decision aid, multiple-choice questions about	Risk perception

(*Continued*)

Table 5. Continued

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
		four of the risks associated with tamoxifen: endometrial cancer, hormonal symptoms, blood clotting, and cataracts; participants were asked to identify which of the following groups was most likely to experience each of these risks: women who take tamoxifen, women who do not take tamoxifen, both groups are equally likely, or don't know)	
Henkemans et al. (2008)	A diabetes self-care computer program in which a computer assistant interacts with the patient	Diabetes knowledge test (containing eight multiple-choice questions dealing with aspects of Type II diabetes)	Experienced usability (concerning effectiveness, efficiency, and satisfaction) Preference
Ubel et al. (2010)	Web-based decision aid with information about tamoxifen (a medication prescribed to women with breast cancer)	Knowledge was assessed with six multiple-choice questions about the risks and benefits of tamoxifen (participants indicated who was more likely to experience each risk and benefit: women who take tamoxifen, women who do not take tamoxifen, both groups are equally likely, or don't know)	Subjective perceptions of risks and benefits
McKenna & Scott (2007)	Leaflets about four topics: occupational theory, arthritis, energy-saving techniques, and stress management	Knowledge test (10 true/false statements about key facts in the leaflets)	Certainty Preference
Makoul et al. (2009)	5-minute multimedia program about colorectal cancer screening	Screening relevant knowledge (10 open-ended knowledge items about anatomy relevant to colorectal cancer and screening age and personal susceptibility and screening tests)	Willingness to consider screening options Ratings of the multimedia program Intention to discuss screening with the physician
Mazor et al. (2007)	Educational health video in which a physician is talking about anticoagulant medication management	Warfarin-related knowledge (a 22-item test of warfarin-related knowledge with closed response options: "true," "false," and "don't know")	Beliefs about warfarin Beliefs about warfarin regimen Beliefs about laboratory regimen Adherence
Walker et al. (2007)	Providing a pictorial mind map to an arthritis booklet	Knowledge: Knowledge Scale Questionnaire, adapted from an existing rheumatoid arthritis knowledge questionnaire for use in clinical settings; the eight sections comprised 40 true/false statements)	
Multiple-choice or open-ended questions for explicit or implicit information			
Shukla et al. (2012)	Informed consent descriptions of cataract surgery	Multiple-choice questions for explicit info (pertaining to specific risks of, benefits of, and treatment alternatives to cataract surgery)	

(*Continued*)

Table 5. Continued

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
Donelle et al. (2009)	Cancer risk information	Comprehension of cancer risk information as measured by a six-item questionnaire: open-ended questions for explicit info (number recognition) and questions that required simple number calculations/operations (implicit info)	
Knapp et al. (2005) (Study 2)	Medication instructions	Interpretation of pictograms (open-ended questions about what instruction or warning the pictogram represented; answers were recorded as correct or incorrect)	
Morrow et al. (2005)	Medication instructions for familiar and unfamiliar medicine	Comprehension (open book), measured with 12 open-ended questions about information that was explicitly stated (e.g., medication name, purpose, dose, times to take, and potential side effects) or implied by the instruction (e.g., how many pills to take in a 24-hour period, what to do if a dose is missed)	Instruction recall (free and cued) Comprehension time
Krieger et al. (2010)	Three messages explaining randomization in Phase III clinical trials	Randomization comprehension measured with three Likert-type items	Attention Message-induced affect Yielding (opinions on explanations) Postintervention behavioral intention to participate in clinical trials
Liu, Kemper, & Bovaird (2009)	Health texts on different topics	Comprehension of texts measured with six yes/no comprehension questions; each question explicitly tested factual information that was directly stated in the text	Subjective difficulty rating
Donelle et al. (2008)	Internet articles with consumer-oriented colorectal cancer prevention information	Comprehension of risk information about colorectal cancer: (a) total risk comprehension scores, (b) risk comprehension scores from the common Internet article, and (c) risk comprehension scores from the uncommon Internet article	
LaVallie et al. (2012)	Risk information for a hypothetical disease and the possible benefits of two different treatments	Comprehension: responses to the three risk comprehension items, coded as correct or incorrect (summary variable reflecting the total number of correct risk comprehension questions for each person)	
Liu, Kemper, & McDowd (2009) (subgroup	Health-related texts on diverse subjects	Text comprehension task (three yes/no comprehension questions were created for each text; these questions did not require	Eye-tracking measures Word recognition task Response times

(*Continued*)

Table 5. Continued

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
of younger and older adults)		inferences, as the answers had been directly stated in the texts) Picture comprehension task (one forced-choice question was created for each text)	
Task performance			
Cardarelli et al. (2011)	Medication bottle labels	Medication identification task: participants' ability to accurately match medication bottles with conditions when placed in front of participants and then at a distance of 2 feet	
Xie (2011)	Online tutorial evaluating Internet health information	e-health literacy efficacy e-health literacy skills	e-health literacy supplemental measures: (a) perceived usefulness of the Internet in helping make health decisions and (b) perceived importance of being able to access health resources on the Internet Attitudes toward the intervention
Bailey et al. (2012)	Prescription (Rx) medication instructions	Rx understanding (demonstration of correct dose, frequency, and spacing)	Regimen dosing ability (demonstration of correct dose, frequency and spacing for five drug regimens) Regimen consolidation (number of times participant would take medication with five drug regimens)
Morrow et al. (2008) (Study 1)	Providing an external aid to medication management	Problem-solving accuracy: accuracy was measured by the total points (out of 24 or 29, depending on specific medications used) awarded for meeting medication requirements	Completion time Efficiency (created by dividing solution time by accuracy, indicating time needed to achieve the same level of accuracy across participants) Subjective workload: NASA-TLX measure composed of 5-point Likert scales that measured mental demand, time pressure, mental effort required, assessed performance, and frustration
Morrow et al. (2008) (Study 2)	Providing an external aid to medication management	Problem-solving accuracy: accuracy was measured by the total points (out of 24 or 29, depending on specific medications used) awarded for meeting medication requirements	Completion time Efficiency (created by dividing solution time by accuracy, indicating time needed to achieve the same level of accuracy across participants) Subjective workload: NASA-TLX measure composed of 5-point Likert scales that measured mental demand, time pressure, mental effort required, assessed performance, and frustration
Henkemans et al. (2008)	A diabetes self-care computer program in which a computer assistant interacts with the patient	Effectiveness (measured by logging the errors made while completing the scenarios) Efficiency (measured by logging the	Experienced usability (concerning effectiveness, efficiency, and satisfaction) Preference

(*Continued*)

Table 5. Continued

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
		time required to fulfill the scenarios and mental effort experienced)	
Recall			
Ruiz et al. (2013)	Multimedia computer-based tutorial about cardiovascular risk	Risk understanding measured by gist and verbatim knowledge (understanding one's own risk and precise numerical representation of events) Risk understanding measured by raw data question and by frequency and percent change question Measured at Time 1 as risk understanding, measured at Time 2 as short-term recall, measured at Time 3 as long-term recall	Confidence Perception of importance Perception of seriousness Intent to adhere to risk factor modification Self-efficacy in performing risk factor modification Actual adherence to risk factor modification Accessibility of the information Attitudes toward the computer program
Freed et al. (2013)	Health information text on colorectal cancer screening	Recognition memory: patients were asked to indicate whether a sentence in the test was old or new (correctly identified old statements were recorded as hits [measure of sensitivity], whereas new statements incorrectly identified as old were recorded as false alarms [measure of response bias])	
Astley et al. (2008)	Information about coronary angiography, delivered as part of the informed consent procedure	Recall of risk information (measured by a 5-point investigator-developed questionnaire)	Satisfaction with the informed consent process Level of anxiety created by disclosure of risk information
Kreuter et al. (2010)	Video with 11 key messages about breast cancer risk	Unprompted recall, coded for (a) any valid response, (b) specific mention of breast cancer or mammography, (c) specific mention of women in the video, (d) specific mention of any video topic or key message	Liking Novelty Learning new information Barriers to mammography Perceived risk Perceived social norms Intention to get mammogram
Van Weert et al. (2011) (subgroup of younger and older adults)	Personalized website with information on surgeries for treating lung cancer	Information recall (measured using an adapted version of the Netherlands Patient Information Recall Questionnaire that consisted of a set of five open-ended questions)	Perceived understandability Time spent on the website Satisfaction
Cloze test			
Todd & Hoffman-Goetz (2011)	Colon cancer information sheet and accompanying cloze test	Comprehension measured by cloze test of colon cancer information sheet	
Griffin et al. (2006)	Educational brochures on various health subjects	Cloze test	
Friedman & Hoffman-Goetz (2007)	Breast, prostate, and colorectal cancer information from the Web	Cloze test	

(*Continued*)

Table 5. Continued

	Health-related document studied	Primary outcome measure comprehension	Secondary outcome measures
Composite measure			
Mittal et al. (2007)	Informed consent form	Understanding score (using a modified MacArthur Competence Assessment Tool for Clinical Research consisting of four subscale scores: (a) understanding relevant information (range = 0–26), (b) appreciation of the applicability/ significance of the information for one's own situation (0–6), (c) reasoning with the information (0–8), and (d) expression of a choice (0–2)	Administration time
Paris et al. (2010)	Informed consent document	Comprehension measured by the Questionnaire d'Evaluation de la Compréhension de l'information Ecrite chez des Malades, consisting of six different domains Objective comprehension (28 questions), subjective comprehension (12 questions: what the participants think they understood)	
Other measures			
Zamarian et al. (2010)	Information about outcomes of 20 unknown medications, presented on a computer screen	Framing effects (computed as score differences between complementary conditions, i.e., between the positive frame–high% condition and the negative frame–low% condition (framing effect-1) and between the positive frame–low% condition and the negative frame–high% condition (framing effect-2))	
Brooke et al. (2013)	Balance appointment leaflets	Finding information (''yes,'' ''no,'' or ''found with difficulty'') Guided reproduction (being able to express the information in one's own words)	
Wilson & Park (2008) (subgroup of younger and older adults)	Health-related statements	Recognition test (participants had to decide whether each statement shown was an unchanged version of a previously studied statement, a changed version of a studied statement, or entirely new)	

of graphical formats (Donelle, Hoffman-Goetz, Gatobu, & Arocha, 2009, MQ; Ruiz et al., 2013, HQ; Zikmund-Fisher, Fagerlin, Roberts, Derry, and Ubel, 2008, FQ). Only in the study of Zikmund-Fisher and colleagues (2008), a limited effect of pictographs compared to other graphical formats was found. In the study of Ruiz and colleagues (2013), recall decreased when risk information was presented in numerical formats with icon arrays compared to numerical formats without icon arrays; no differences were found in understanding. Donelle and colleagues (2009) found no differences in understanding between text-only and graphical formats. Thus, the evidence

Table 6. Results of the quality assessment by study

Source	Quality score[a]	Reporting	External validity	Internal validity		
				Bias	Confounding	Construct validity/reliability
Older adults						
Ruiz et al. (2013)	20	1,2,3,4,5,6,7,8	9,10 (11)	12,14,15 (13)	16,17,18	19,20,21
Kreuter et al. (2010)	20	1,2,3,4,5,6,7,8	9,10,11	12,14,15 (13)	16,17,18	19,20,21
Walker et al. (2007)	20	1,2,3,4,5,6,7,8	9,10,11	12,14,15 (13)	16,17,18	19,20,21
Mazor et al. (2007)	20	1,2,3,4,5,6,7,8	9,10,11	12,14,15 (13)	16,17,18	19,20,21
Gattellari & Ward (2005)	19	1,2,3,4,5,6,7,8	9,10,11	12,14,15 (13)	16,17,18	19,20 (21)
Volk et al. (2008)	19	1,2,3,4,5,6,7,8	9,10,11	12,15 (13,14)	16,17,18	19,20,21
Freed et al. (2013)	18	1,2,3,4,5,6,8 (7)	9,10 (11)	12,14,15 (13)	16,17,18	19,20,21
Xie (2011)	18	1,2,3,4,5,6,7,8	11 (9,10)	12,14,15 (13)	16,17,18	19,20,21
Morrow et al. (2008) (1)	18	1,2,3,4,5,6,7,8	9 (10,11)	12,14,15 (13)	16,17,18	19,20,21
Morrow et al. (2008) (2)	18	1,2,3,4,5,6,7,8	9 (10,11)	12,14,15 (13)	16,17,18	19,20,21
Ilic et al. (2008)	18	1,2,3,4,5,6,7,8	10,11 (9)	12,14,15 (13)	16,17 (18)	19,20,21
Sudore et al. (2007)	18	1,2,3,4,5,6,8 (7)	9,10,11	12,15 (13,14)	16,17,18	19,20,21
Mittal et al. (2007)	18	1,2,3,4,5,7,8 (6)	9,10 (11)	12,14,15 (13)	16,17,18	19,20,21
Bailey et al. (2012)	17	1,2,3,5,6,7 (4,8)	9,10,11	12,14,15 (13)	16,17,18	19,20 (21)
Todd & Hoffman-Goetz (2011)	17	1,2,3,4,6,7 (5,8)	9,10,11	12,15 (13,14)	16,17,18	19,20,21
Makoul et al. (2009)	17	1,2,3,4,6,7,8 (5)	9,10,11	12,14 (13,15)	16,17,18	20,21 (19)
Zikmund-Fisher et al. (2008)	17	1,2,3,4,5,8 (6,7)	10,11 (9)	12,14,15 (13)	16,17,18	19,20,21
Friedman & Hoffman-Goetz (2007)	17	1,2,3,4,6,7,8 (5)	10 (9,11)	12,14,15 (13)	16,17,18	19,20,21
McKenna & Scott, 2007)	17	1,2,3,4,6,7,8 (5)	9,10,11	12,14,15 (13)	16 (17,18)	19,20,21
Shukla et al. (2012)	16	1,2,3,5,6,7,8 (4)	9,10,11	12 (13,14,15)	16,17,18	20,21 (19)
Krieger et al. (2010)	16	1,2,3,4,6,7,8 (5)	9,11 (10)	12,14,15 (13)	16,17 (18)	19,20,21
Frosch et al. (2008)	16	1,2,3,4,5,6,7 (8)	9,10,11	12,14,15 (13)	16,18 (17)	21 (19,20)
Astley et al. (2008)	16	1,2,3,4,5,6,7,8	9,10,11	12,15 (13,14)	16,17 (18)	21 (19,20)
Griffin et al. (2006)	16	1,2,3,4,6,7,8 (5)	9,10 (11)	12,14,15 (13)	18 (16,17)	19,20,21
Morrow et al. (2005)	16	1,2,3,4,5,6,7,8	10 (9,11)	12,14,15 (13)	16,18 (17)	20,21 (19)
Liu, Kemper, & Bovaird (2009)	15	1,2,3,4,6,7,8 (5)	10 (9,11)	12,14,15 (13)	18 (16,17)	19,20,21
Paris et al. (2010)	15	1,2,3,5,6,7 (4,8)	9,10 (11)	12,15 (13,14)	16,17 (18)	19,20,21
Ubel et al. (2010)	14	1,2,3,4,6,7 (5,8)	10 (9,11)	12,14,15 (13)	16,17 (18)	19,21 (20)
Zamarian et al. (2010)	14	1,2,3,5,7,8 (4,6)	10 (9,11)	12,14,15 (13)	18 (16,17)	20,21 (19)
Donelle et al. (2009)	14	1,2,3,4,6,7,8 (5)	9,10 (11)	12,15 (13,14)	16,17 (18)	21 (19,20)
Knapp et al. (2005) (2)	14	1,2,4,6,8 (3,5,7)	9,10 (11)	12,14,15 (13)	16,18 (17)	20,21 (19)
Cardarelli et al. (2011)	13	1,2,3,4,6,7 (5,8)	9,10 (11)	12,15 (13,14)	(16,17,18)	19,20,21
LaVallie et al. (2012)	12	1,2,3,4,6,7,8 (5)	(9,10,11)	12,15 (13,14)	16,18 (17)	21 (19,20)
Donelle et al. (2008)	12	1,2,3,6,7 (4,5,8)	10 (9,11)	12,15 (13,14)	18 (16,17)	19,20,21
Henkemans et al. (2008)	11	1,2,4,7,8 (3,5,6)	11 (9,10)	12,14,15 (13)	18 (16,17)	19 (20,21)
Brooke et al. (2013)	9	1,2,3,6 (4,5,7,8)	10 (9,11)	15 (12,13,14)	16 (17,18)	20,21 (19)
Younger vs. older subgroups						
Liu, Kemper, & McDowd (2009)	16	1,2,3,4,5,6,7 (8)	9,10 (11)	12,14,15 (13)	16,17 (18)	20,21 (19)
Van Weert et al. (2011)	14	1,2,3,4,6,7,8 (5)	(9,10,11)	12,14 (13,15)	16,17 (18)	19,20,21
Wilson & Park (2008)	12	1,2,4,6,7 (3,5,8)	10 (9,11)	12,14,15 (13)	(16,17,18)	19,20,21

Note. Items in parentheses were regarded as not sufficient. See Table 2 for definitions of criteria numbers.
[a]Validity scores were calculated as the sums of all items. Each item was given equal weight.

regarding the effectiveness of graphical formats in improving comprehensibility is inconsistent.

Four studies assessed the added value of pictures (Cardarelli et al., 2011, MQ; Knapp, Raynor, Jebar, & Price, 2005, MQ; Liu, Kemper, & McDowd, 2009, FQ; Van Weert et al., 2011, MQ). Three studies compared texts with and without pictures. Two studies found effects of adding pictures on the correct identification of medications (Cardarelli et al., 2011) and recall and perceived understandability of a website (Van Weert et al., 2011). However, Liu, Kemper, and McDowd (2009) found no differences in comprehension between texts with and without illustrations. Finally, Knapp and colleagues (2005) compared pictures of different sizes. All in all, the level of evidence regarding the effectiveness of pictures for improving comprehensibility is inconsistent.

One study focused on another textual design feature, namely, ordering principles combined with extra (context)

information in medication risk information. Ubel and colleagues (2010, MQ) showed that the order in which information is presented matters, but only if no context information about competing health risks is provided. There is inconclusive evidence for a possible influence of the order in which information is presented.

Eight studies assessed the possible effects of changing different linguistic characteristics, that is, simplifying language, or changing other linguistic characteristics (e.g., translation into one's first language). Five studies focused on the effects of simplifying language (sometimes assessed with a grade-based readability formula) on comprehensibility and examined whether presenting readers with documents that differed in readability according to the outcomes of formulas like the Flesch reading ease (RE; Kincaid, Fishburne, Rogers, & Chissom, 1975) actually resulted in different scores for comprehensibility. Three of these studies specifically examined whether simplifying language improved comprehensibility. Two studies found higher comprehension for the simplified versions (Shukla et al., 2012, FQ; Van Weert et al., 2011, MQ), whereas one study found no differences (Paris et al., 2010, FQ). Two studies examined whether original documents that differed in readability scores as calculated with the Flesch RE formula (Liu, Kemper, & Bovaird, 2009, FQ) or with the Simple Measure of Gobbledygook (SMOG) formula (Friedman & Hoffman-Goetz, 2007, FQ) also differed in comprehensibility. Liu, Kemper, and Bovaird (2009, FQ) found that improved readability according to Flesch RE did not affect comprehension for older adults with larger working memories, but older adults with smaller working memories had even more difficulty understanding the texts with higher readability scores according to Flesch RE.[1] In the study of Friedman and Hoffman-Goetz (2007, FQ) on comprehension of Web texts with information on three topics, differences in comprehension scores between texts with different complexity levels were only significant for information on one of the topics. All in all, evidence for the effectiveness of simplifying language is inconsistent. Furthermore, it remains unclear how exactly in these five studies the intended linguistic simplification of health-related documents was achieved.

Three studies examined the effectiveness of other linguistic characteristics: the effect of different types of metaphors for explaining the concept of randomization in clinical cancer trials (Krieger, Parrott, & Nussbaum, 2010, FQ), the effect of an information sheet in Chinese immigrant women's first (Chinese) or second (English) language (Todd & Hoffman-Goetz, 2011, FQ), and the difference in comprehension between common (familiar) and uncommon (unfamiliar) cancer prevention information[2] (Donelle, Arocha, & Hoffman-Goetz, 2008, MQ). No differences in comprehension were found among the three randomization messages (Krieger et al., 2010). Chinese immigrant women performed significantly better with information offered in their first language compared to their second language (Todd & Hoffman-Goetz, 2011). Donelle and colleagues (2008) found better comprehension for common cancer prevention information. Taken together, the evidence for effective interventions on linguistic characteristics of health-related documents to improve comprehensibility is inconsistent.

Six studies assessed the possible effects of multiple-feature revisions. Five studies reported on multiple-feature document revisions that included revisions of textual design and linguistic characteristics (Brooke, Herbert, Isherwood, Knapp, & Raynor, 2013, PQ; Freed et al., 2013, HQ; McKenna & Scott, 2007, FQ; Morrow et al., 2005, FQ; Sudore et al., 2007, HQ). Two studies found evidence for the effectiveness of such a revision (Freed et al., 2013; McKenna & Scott, 2007), two studies found mixed results (Brooke et al., 2013; Morrow et al., 2005), and one study found no differences in comprehension between the original and revised documents (Sudore et al., 2007). The final study, by Bailey, Sarkar, Chen, Schillinger, and Wolf (2012, FQ), studied the effects of medication instructions in which health literacy best practices were followed (e.g., grounding medication-taking time to four distinct time periods, using simpler terms). These researchers found higher comprehension for instructions with these multiple-feature revisions of linguistic characteristics. Taken together, these studies provide weak evidence for the effectiveness of multiple-feature revisions based on content and design principles.

We identified five studies that focused on the design of informational content of health-related documents by examining the effects of different ways of framing the message or the effects of narrative formats. *Framing* refers to the way messages are worded without changing the content and includes the use of positive or negative language and also the use of gain or loss frames, in which the focus lies either on the positive effects of behaving according to the advice in the message (gain frame) or on the negative effects of not adhering to the advice (loss frame). Three studies compared the effects of positive and negative framing of health information. Two studies found that positive or negative language can influence comprehension (Wilson & Park, 2008, PQ; Zamarian, Benke, Buchler, Wenter, & Delazer, 2010, MQ), whereas one study found no differences in knowledge increases between a positively framed and a negatively framed text (Makoul et al., 2009, FQ). Together, these studies provide inconclusive evidence of effects on comprehension.

Two studies assessed the effectiveness of using narrative formats. Kreuter and colleagues (2010, HQ) presented African American women with a video with 11 key messages about breast cancer risk in informational versus narrative format. Their narrative video was more effective with regard to the recall of relevant information 3 and 6 months after watching the video. Mazor and colleagues (2007, HQ) presented middle-aged and older patients on anticoagulation

[1] The authors explained this counterintuitive finding by stating that texts with higher readability scores according to Flesch RE (shorter words and sentences) may have suffered from decreased textual cohesion, making it more difficult for older adults with smaller working memories to integrate the information content.

[2] "Common information included material that was widely publicized, easily accessible, and that replicated general CCS [Canadian Cancer Society] introductory information available for all cancer types" (Donelle et al., 2008, p. 2).

medication with a video in three different formats: narrative information, statistical information, or a combination of both narrative and statistical information. Patients who viewed a version with narrative information (either narrative only or narrative combined with statistical information) showed greater knowledge gains compared to patients who viewed a version with only statistical information. Taken together these studies provide moderate evidence for the added value of a narrative format.

Five studies examined other features of health-related documents or context factors that were difficult to classify. Xie (2011, HQ) studied the effect of different information presentation channels of an online tutorial. In addition, she looked at whether older adults benefited more from a collaborative learning method compared to an individualistic learning method. No differences in effectiveness were found. Henkemans and colleagues (2008, PQ) assessed the effectiveness of two different types of diabetes self-care computer assistants. They did not find differences in knowledge increase between the two conditions. With regard to task performance (speed and accuracy), participants using the adaptive assistant were faster and made less errors. The level of evidence for the effectiveness of both learning method and type of computer assistant is inconclusive. Zikmund-Fisher and colleagues (2008, FQ) looked at different types of descriptions of medication risk, and LaVallie, Wolf, Jacobsen, Sprague, and Buchwald (2012, MQ) studied different treatment benefit descriptions. They found differences in comprehension for different descriptions of risk and benefits, but because of the differences between descriptions in these two studies, the level of evidence for the effectiveness of descriptions of numerical information is inconclusive (see Table 5).

Two studies focused on the effectiveness of providing an external aid for reading and understanding health-related information. Morrow and colleagues (2008, HQ) found that their older participants were more accurate and efficient when they used an external aid in a role-play task on medication management. Walker and colleagues (2007, HQ) found no significant difference in knowledge increase between participants with rheumatoid arthritis who received a booklet with or without a pictorial mind map. All in all, the level of evidence for the effectiveness of external aids is inconsistent.

Health Literacy Subgroups

Eighteen out of 38 studies examined whether there was a main effect of (health) literacy, numeracy, education level, or cognitive measures on comprehension of health-related documents. Thirteen of these 18 studies reported that lower scores on (health) literacy, numeracy, education, or cognitive abilities were associated with lower levels of comprehension of health-related documents. In the remaining five studies, possible differences in comprehension scores did not reach significance. Six studies also examined whether there was an interaction between level of health literacy (or other cognitive measures) and type of intervention on the comprehension scores of older adults (Bailey et al., 2012; Gattellari & Ward, 2005; Liu, Kemper, & Bovaird, 2009; Paris et al., 2010; Volk et al., 2008; Zamarian et al., 2010). Only two studies found such an interaction effect, for cognitive measures in general (Zamarian et al., 2010) and for working memory and verbal ability (Liu, Kemper, & Bovaird, 2009). Liu, Kemper, and Bovaird (2009) found that readability levels of health texts measured by Flesch RE had no effect on comprehension for older adults with larger working memories, although older adults with smaller working memories had more difficulty understanding texts in which readability measured by Flesch RE was lower. Zamarian and colleagues (2010) found stronger framing effects for older adults with poorer cognitive performance.

Discussion

The present study is the first to systematically review the evidence for the effectiveness of interventions aimed at improving the comprehensibility of health-related documents for older adults, paying special attention to effects of health literacy. Unlike earlier reviews in the general population (e.g., Berkman et al., 2011; Sheridan et al., 2011), we did not find consistent evidence for the effectiveness of interventions manipulating features and formats of health-related documents that aim to enhance comprehensibility in older adults. However, two sets of interventions were identified that seem to hold some promise for enhancing the comprehensibility of health-related documents for older adults: multiple-feature revisions and the use of narrative formats.

The studies in this review provide weak evidence for the effectiveness of such multiple-feature revisions. However, it is hard to draw specific conclusions on the particular features that may contribute to comprehensibility, precisely because of the multiplicity of features targeted in these revisions. Moreover, because the application of recommendations found in the health literacy literature mostly results in such multiple-feature revisions, it is remarkable that the studies in our review fall short of providing strong evidence for these interventions.

This review provides moderate evidence of the benefits of using a narrative format in health-related documents, with the two studies that we found on narratives (Kreuter et al., 2010; Mazor et al., 2007) both reporting benefits for narrative video formats. These positive results are confirmed in studies inside as well as outside the field of health communication (e.g., Graesser & Ottati, 1996; Thompson & Kreuter, 2013). The effectiveness of narratives in enhancing the comprehensibility of health-related documents may be due to their recognizable format. Narrative communication is a mode of interaction people use frequently in their daily lives. Story structures are very familiar and may therefore be easier to process compared to the structure of a less familiar type of text, such as a patient information leaflet. Narrative forms of health-related documents may hence place lower demands on processing capacities and be easier to process (Hinyard & Kreuter, 2007). Furthermore, interventions that include narrative formats may increase personal involvement and thereby enhance motivation and engagement for processing health-related documents (Hinyard & Kreuter, 2007). It is important to note that both narrative interventions included

video formats, which introduces the possibility that the positive effect of these interventions was due to the specific combination of video and narrative. The effects of using only narratives in this target group thus deserve additional study.

One other point to note is that we found some weak evidence for the ineffectiveness of multimedia formats in health documents compared to a single media format. This finding highlights the importance of theories of multimedia understanding, which posit that understanding multimedia requires integrative information processing from different sensory modalities (Mayer, 2005). Older adults may have trouble with this kind of processing, as both a decline in processing capacity and sensory deficits presumably influence the integration of information. Evidently it is necessary to incorporate instructional theories such as cognitive load theory and cognitive theory of multimedia learning in designing optimal multimedia health-related documents for older adults. This may be accomplished by mapping "age-related cognitive declines on the potentially compensatory strategies offered by existing instructional theories" (Van Gerven, Paas, & Tabbers, 2006, p. 149; see also Paas, Van Gerven, & Tabbers, 2005; Watkins & Xie, 2014; Wolfson & Kraiger, 2014). Only two out of six studies on multimedia in this review referred to multimedia theories. Mittal and colleagues (2007) implicitly referred to the modality effect when they expected greater understanding for "simultaneous presentation of visual and verbal information," including voice narration and video, whereas Xie (2011) explicitly referred to the redundancy effect when she hypothesized a learning decrease when "identical information is presented in multiple media forms."

Evidence for the effectiveness of all other interventions turned out to be either inconsistent or inconclusive. This finding seems to stand in contrast to earlier reviews (e.g., Sheridan et al., 2011) that did find consistent evidence for interventions such as the revision of several discrete design features. There are at least three possible reasons for this discrepancy. First, the more recent studies that were sampled in the current review may simply have been less positive about the various interventions. Second, the studies that were included in our review specifically focused on older adults, which is different from the more general review by Sheridan and colleagues (2011). In addition, part of the studies in our review used different interventions from those included in the Sheridan and colleagues review. For instance, Sheridan and colleagues did not include studies on framing, narrative versus informational formats, and external aids. And third, in our current systematic review, we applied a rather strict analysis of the levels of evidence for a given intervention, in which we only considered evidence consistent and sufficient if multiple fair- to high-quality studies showed proof for the effectiveness of design features and formats. In the Sheridan and colleagues review, it was relatively often the case that evidence was regarded as positive if a single study—and thus with only a single operationalization of a given factor—showed a positive effect. However, Sheridan and colleagues also concluded that it is precisely the heterogeneity of studies in terms of operationalization that hinders the generalizability of findings and strength of evidence. We agree with this conclusion, and thus we would like to stress the importance of (a) using standardized research methodologies (e.g., standardized and sensitive measures of comprehension) and (b) conducting replications in intervention studies on health-related documents. Furthermore, the limited and inconsistent evidence in this review shows that designers of health-related documents for older adults should be on their guard when they apply general design principles, whether or not it is claimed that these principles were tested in empirical studies in the target group. Finally, this shows the importance of the role of theory in guiding intervention studies that are appropriate for the target group of older adults (e.g., integrating theories on cognitive aging and multimedia and text comprehension into intervention development; see, e.g., Johnson, 2003).

Only a relatively small proportion of studies explicitly examined whether the effectiveness of interventions differed between participants with different levels of health literacy, education, or cognitive capabilities. Additional research is thus needed to determine which measures are effective at improving health documents, especially in elderly people with poor health literacy.

Strengths and Limitations

This is the first systematic review focusing on the effectiveness of interventions aiming to improve the comprehensibility of health-related documents in older adults. Strengths of this review are (a) the broad range of intervention topics that were investigated, (b) the broad selection of comprehensibility outcome measures that were considered, (c) the focus on the older population, and (d) the strict levels of evidence assessment by multiple reviewers. It should be noted that our review also has limitations. First, only studies from 2005 onward were included, which might have led us to miss important earlier studies. However, research on this topic has gained interest relatively recently, making this rather unlikely. Moreover, we did a quick scan of $<$2005 references of the included studies. Articles found from these references did not alter our conclusions. Second, we were unable to make a quantitative summary of effects, which makes it difficult to get a complete and objective overview of effects.

Implications

Because of their specific characteristics, the use of narrative formats seems to be a promising strategy, especially for older adults with limited levels of health literacy. Because our review included only two studies on the effects of narratives in health communication targeted at older people, more research is warranted to determine whether and how the use of narrative formats in health-related documents may contribute to comprehensibility, especially in older adults with limited health literacy.

Furthermore, research should further explore the effectiveness of multiple-feature revisions for older adults. Although these interventions are widespread, our review

only provided weak evidence for the effectiveness of these interventions in older adults. Therefore, more systematic research is needed to determine exactly which features contribute to comprehensibility and which effects may be expected in which combinations with other features.

Our results underline that practitioners and designers of health-related documents for older adults should be cautious in applying interventions that have proven successful in general populations, because these will not necessarily be beneficial for older adults.

Conclusion

Our review shows that evidence for interventions aiming to improve the comprehensibility of health-related documents for older adults with different levels of health literacy is limited. However, narratives and multiple-feature revisions are promising. These may provide good routes to improving the health of older adults. In addition, more research is needed to analyze the effectiveness of their separate components, especially for this particular population.

Acknowledgments

We are very thankful to research assistants Aline Douma, Mayke Musters, Jacqueline Quartero, and Bram Roenhorst for their contributions to this article. Furthermore, we are very grateful to Monika Koester of BZgA: Bundeszentrale für gesundheitliche Aufklärung, who contributed to the pilot session for the title review.

Funding

This research was supported under the Intervention Research on Health Literacy Among Ageing Population (IROHLA) project. IROHLA is coordinated by the University Medical Center Groningen and has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 305831.

Supplemental Material

A supplemental appendix to this article (Table A1: Summary of the included intervention studies on health related documents, including the effect of health literacy) is available online at http://dx.doi.org/10.1080/10810730.2015.1049306.

References

Abraham, C., & Kools, M. (Eds.). (2012). *Writing health communication: An evidence-based guide*. London, UK: Sage.

Adams, R. J., Piantadosi, C., Ettridge, K., Miller, C., Wilson, C., Tucker, G., & Hill, C. L. (2013). Functional health literacy mediates the relationship between socio-economic status, perceptions and lifestyle behaviors related to cancer risk in an Australian population. *Patient Education and Counseling*, *91*, 206–212. doi: 10.1016/j.pec.2012.12.001

Al Sayah, F., Majumdar, S. R., Williams, B., Robertson, S., & Johnson, J. A. (2013). Health literacy and health outcomes in diabetes: A systematic review. *Journal of General Internal Medicine*, *28*, 444–452. doi: 10.1007/s11606–012-2241-z

Ashida, S., Goodman, M., Pandya, C., Koehly, L. M., Lachance, C., Stafford, J., & Kaphingst, K. A. (2011). Age differences in genetic knowledge, health literacy and causal beliefs for health conditions. *Public Health Genomics*, *14*, 307–316. doi: 10.1159/000316234

Astley, C. A., Chew, D. P., Aylward, P. E., Molloy, D. A., & De Pasquale, C. G. (2008). A randomised study of three different informational AIDS prior to coronary angiography, measuring patient recall, satisfaction and anxiety. *Heart, Lung and Circulation*, *17*, 25–32. doi: 10.1016/j.hlc.2007.04.008

Bailey, S. C., Sarkar, U., Chen, A. H., Schillinger, D., & Wolf, M. S. (2012). Evaluation of language concordant, patient-centered drug label instructions. *Journal of General Internal Medicine*, *27*, 1707–1713. doi: 10.1007/s11606–012-2035–3

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., Viera, A., Crotty, K., … Viswanathan, M.. (2011). *Health literacy interventions and outcomes: An updated systematic review* (Evidence Reports/Technology Assessments, No. 199). Rockville, MD: Agency for Healthcare Research and Quality.

Bostock, S., & Steptoe, A. (2012). Association between low functional health literacy and mortality in older adults: Longitudinal cohort study. *British Medical Journal*, *344*, e1602–e1602. doi: 10.1136/bmj.e1602

Brooke, R. E., Herbert, N. C., Isherwood, S., Knapp, P., & Raynor, D. K. (2013). Balance appointment information leaflets: Employing performance-based user-testing to improve understanding. *International Journal of Audiology*, *52*(3), 162–168. doi: 10.3109/14992027.2012.745200

Cardarelli, R., Mann, C., Fulda, K. G., Balyakina, E., Espinoza, A., & Lurie, S. (2011). Improving accuracy of medication identification in an older population using a medication bottle color symbol label system. *BMC Family Practice*, *12*, 142. doi: 10.1186/1471-2296-12-142

Chin, J., Morrow, D. G., Stine-Morrow, E. A. L., Conner-Garcia, T., Graumlich, J. F., & Murray, M. D. (2011). The process-knowledge model of health literacy: Evidence from a componential analysis of two commonly used measures. *Journal of Health Communication*, *16*(Suppl. 3), 222–241. doi:10.1080/10810730.2011.604702

Coleman, C. A., Hudson, S., & Maine, L. (2013). Health literacy practices and educational competencies for health professionals: A consensus study. *Journal of Health Communication*, *18*(Suppl. 1), 82–102. doi:10.1080/10810730.2013.829538

Creemers, H. E., Verhulst, F. C., & Huizink, A. C. (2009). Temperamental risk factors for adolescent cannabis use: A systematic review of prospective general population studies. *Substance Use & Misuse*, *44*, 1833–1854. doi: 10.3109/10826080802494933

Donelle, L., Arocha, J. F., & Hoffman-Goetz, L. (2008). Health literacy and numeracy: Key factors in cancer risk comprehension. *Chronic Diseases in Canada*, *29*(1), 1–8.

Donelle, L., Hoffman-Goetz, L., Gatobu, S., & Arocha, J. F. (2009). Comprehension of Internet-based numeric cancer information by older adults. *Informatics for Health and Social Care*, *34*(4), 209–224. doi: 10.3109/17538150903358552

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*, 377–384. doi: 10.1136/jech.52.6.377

Freed, E., Long, D., Rodriguez, T., Franks, P., Kravitz, R. L., & Jerant, A. (2013). The effects of two health information texts on patient recognition memory: A randomized controlled trial. *Patient Education and Counseling*, *92*, 260–265. doi: 10.1016/j.pec.2013.03.008

Friedman, D. B., & Hoffman-Goetz, L. (2007). An exploratory study of older adults' comprehension of printed cancer information: Is readability a key factor? *Journal of Health Communication*, *12*(5), 423–437. doi: 10.1080/10810730701438658

Frosch, D. L., Legare, F., & Mangione, C. M. (2008). Using decision aids in community-based primary care: A theory-driven evaluation with ethnically diverse patients. *Patient Education and Counseling*, *73*(3), 490–496. doi: 10.1016/j.pec.2008.07.040

Gattellari, M., & Ward, J. E. (2005). A community-based randomised controlled trial of three different educational resources for men about prostate cancer screening. *Patient Education and Counseling*, *57*(2), 168–182. doi: 10.1016/j.pec.2004.05.011

Graesser, A. C., & Ottati, V. (1996). Why stories? Some evidence, questions, and challenges. In R. S. Wyer (Ed.), *Knowledge and memory: The real story* (pp. 121–132). Hillsdale, NJ: Erlbaum.

Griffin, J., McKenna, K., & Tooth, L., (2006). Discrepancy between older clients' ability to read and comprehend and the reading level of written educational materials used by occupational therapists. *American Journal of Occupational Therapy*, *60*, 70–80.

Henkemans, O. A. B, Rogers, W. A., Fisk, A. D., Neerincx, M. A., Lindenberg, J., & van der Mast, C. A. P. G. (2008). Usability of an adaptive computer assistant that improves self-care and health literacy of older adults. *Methods of Information in Medicine*, *47*(1), 82–87. doi: 10.3414/me9105

Hinyard, L. J., & Kreuter, M. W. (2007). Using narrative communication as a tool for health behavior change: A conceptual, theoretical, and empirical overview. *Health Education and Behavior*, *34*(5), 777–792. doi: 10.1177/1090198106291963

HLS-EU Consortium. (2012). *Comparative report on health literacy in eight EU member states*. The European Health Literacy Survey (HLS-EU). Retrieved from http://www.maastrichtuniversity.nl/web/Institutes/FHML/CAPHRI/DepartmentsCAPHRI/InternationalHealth/ResearchINTHEALTH/Projects/HealthLiteracyHLSEU/MeasuringHealthLiteracyInEurope.htm

Ilic, D., Egberts, K., McKenzie, J. E., Risbridger, G., & Green, S. (2008). Informing men about prostate cancer screening: A randomized controlled trial of patient education materials. *Journal of General Internal Medicine*, *23*(4), 466–471. doi: 10.1007/s11606-007-0466-z

Johnson, R. E. (2003). Aging and the remembering of text. *Developmental Review*, *23*(3), 261–346. doi: 10.1016/s0273-2297(03)00009-1

Jovic-Vranes, A., & Bjegovic-Mikanovic, V. (2012). Which women patients have better health literacy in Serbia? *Patient Education and Counseling*, *89*, 209–212. doi: 10.1016/j.pec.2012.06.001

Kim, S. H. (2009). Health literacy and functional health status in Korean older adults. *Journal of Clinical Nursing*, *18*, 2337–2343. doi: 10.1111/j.1365-2702.2008.02739.x

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*. Research Branch Report 8–75. Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN.

Knapp, P., Raynor, D. K., Jebar, A. H., & Price, S. J. (2005). Interpretation of medication pictograms by adults in the UK. *Annals of Pharmacotherapy*, *39*, 1227–1233. doi: 10.1345/aph.1e483

Kreuter, M. W., Holmes, K., Alcaraz, K., Kalesan, B., Rath, S., Richert, M., … Clark, E. M. (2010). Comparing narrative and informational videos to increase mammography in low-income African American women. *Patient Education and Counseling*, *81*(Suppl. 1), S6–S14. doi:10.1016/j.pec.2010.09.008

Krieger, J. L., Parrott, R. L., & Nussbaum, J. F. (2010). Metaphor use and health literacy: A pilot study of strategies to explain randomization in cancer clinical trials. *Journal of Health Communication*, *16*(1), 3–16. doi: 10.1080/10810730.2010.529494

Kwan, B., Frankish, J., Rootman, I., Zumbo, B., Kelly, K., Begoray, D., … Hayes, M. (2006). *The development and validation of measures of "health literacy" in different populations*. Vancouver, British Columbia, Canada: University of British Columbia, Institute of Health Promotion Research, and University of Victoria Centre for Community Health Promotion Research.

LaVallie, D. L., Wolf, F. M., Jacobsen, C., Sprague, D., & Buchwald, D. S. (2012). Health numeracy and understanding of risk among older American Indians and Alaska natives. *Journal of Health Communication: International Perspectives*, *17*(3), 294–302. doi: 10.1080/10810730.2011.626497

Liu, C., Kemper, S., & Bovaird, J. A. (2009). Comprehension of health-related written materials by older adults. *Educational Gerontology*, *35*(7), 653–668. doi: 10.1080/03601270902885504

Liu, C., Kemper, S., & McDowd, J. (2009). The use of illustration to improve older adults' comprehension of health-related information: Is it helpful? *Patient Education and Counseling*, *76*(2), 283–288. doi: 10.1016/j.pec.2009.01.013

Makoul, G., Cameron, K. A., Baker, D. W., Francis, L., Scholtens, D., & Wolf, M. S. (2009). A multimedia patient education program on colorectal cancer screening increases knowledge and willingness to consider screening among Hispanic/Latino patients. *Patient Education and Counseling*, *76*(2), 220–226. doi: 10.1016/j.pec.2009.01.006

Mayer, R. E. (2005). *Cognitive theory of multimedia learning: The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mazor, K. M., Baril, J., Dugan, E., Spencer, F., Burgwinkle, P., & Gurwitz, J. H. (2007). Patient education about anticoagulant medication: Is narrative evidence or statistical evidence more effective? *Patient Education and Counseling*, *69*, 145–157. doi: 10.1016/j.pec.2007.08.010

McKenna, K., & Scott, J. (2007). Do written education materials that use content and design principles improve older people's knowledge? *Australian Occupational Therapy Journal*, *54*(2), 103–112. doi: 10.1111/j.1440–1630.2006.00583.x

Mittal, D., Palmer, B. W., Dunn, L. B., Landes, R., Ghormley, C., Beck, C., … Jeste, D. V. (2007). Comparison of two enhanced consent procedures for patients with mild Alzheimer disease or mild cognitive impairment. *American Journal of Geriatric Psychiatry*, *15*(2), 163–167. doi: 10.1097/jgp.0b013e31802dd379

Morrow, D., Clark, D., Tu, W., Wu, J., Weiner, M., Steinley, D., & Murray, M. D. (2006). Correlates of health literacy in patients with chronic heart failure. *The Gerontologist*, *46*(5), 669–676. doi: 10.1093/geront/46.5.669

Morrow, D., Raquel, L., Schriver, A., Redenbo, S., Rozovski, D., & Weiss, G. (2008). External support for collaborative problem solving in a simulated provider/patient medication scheduling task. *Journal of Experimental Psychology: Applied*, *14*(3), 288–297. doi: 10.1037/a0012809

Morrow, D. G., Weiner, M., Young, J., Steinley, D., Deer, M., & Murray, M. D. (2005). Improving medication knowledge among older adults with heart failure: A patient-centered approach to instruction design. *The Gerontologist*, *45*(4), 545–552. doi: 10.1093/geront/45.4.545

Mõttus, R., Johnson, W., Murray, C., Wolf, M. S., Starr, J. M., & Deary, I. J. (2014). Towards understanding the links between health literacy and physical health. *Health Psychology*, *33*, 164–173. doi:10.1037/a0031439

Omachi, T. A., Sarkar, U., Yelin, E. H., Blanc, P., & Katz, P. (2013). Lower health literacy is associated with poorer health status and outcomes in chronic obstructive pulmonary disease. *Journal of General Internal Medicine*, *28*, 74–81. doi: 10.1007/s11606-012-2177-3

Ownby, R. L., Waldrop-Valverde, D., & Taha, J. (2012). Why Is health literacy related to health? An exploration among U. S. national assessment of adult literacy participants 40 years of age and older. *Educational Gerontology*, *38*, 776–787. doi: 10.1080/03601277.2011.645441

Paas, F., Van Gerven, P. W. M., & Tabbers, H. K. (2005). The cognitive aging principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 339–354). New York, NY: Cambridge University Press.

Paris, A., Brandt, C., Cornu, C., Maison, P., Thalamas, C., & Cracowski, J. (2010). Informed consent document improvement does not increase patients' comprehension in biomedical research. *British Journal of Clinical Pharmacology*, *69*(3), 231–237. doi: 10.1111/j.1365-2125.2009.03565.x

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., & Britten, N. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. A product from the ESRC Research Methods Programme. Version 1. Swindon, UK: Economic and Social Research Council.

Ruiz, J. G., Andrade, A. D., Garcia-Retamero, R., Anam, R., Rodriguez, R., & Sharit, J. (2013). Communicating global cardiovascular risk: Are icon arrays better than numerical estimates in improving understanding, recall and perception of risk? *Patient Education and Counseling*, *93*, 394–402. doi: 10.1016/j.pec.2013.06.026

Sheridan, S. L., Halpern, D. J., Viera, A. J., Berkman, N. D., Donahue, K. E., & Crotty, K. (2011). Interventions for individuals with low health literacy: A systematic review. *Journal of Health Communication*, *16*(Suppl. 3), 30–54. doi:10.1080/10810730.2011.604391

Shukla, A. N., Daly, M. K., & Legutko, P. (2012). Informed consent for cataract surgery: Patient understanding of verbal, written, and videotaped information. *Journal of Cataract & Refractive Surgery*, *38*(1), 80–84. doi: 10.1016/j.jcrs.2011.07.030

Smart, K. A., Parker, R. S., Lampert, J., & Sulo, S. (2012). Speaking up: Teens voice their health information needs. *Journal of School Nursing*, *28*(5), 379–388. doi: 10.1177/1059840512450916

Sudore, R. L., Landefeld, C. S., Barnes, D. E., Lindquist, K., Williams, B. A., Brody, R., & Schillinger, D. (2007). An advance directive redesigned to meet the literacy level of most adults: A randomized trial. *Patient Education and Counseling*, *69*, 165–195. doi: 10.1016/j.pec.2007.08.015

Thompson, T., & Kreuter, M. W. (2013). Narrative-based health communication interventions: Using survivor stories to increase breast cancer knowledge and promote mammography. In D. K. Kim, A. Singhal, & G. L. Kreps (Eds.), *Health communication: Vol. 5. Global health communication strategies in the 21st century: Design, implementation, and evaluation* (pp. 118–133). New York, NY: Peter Lang.

Todd, L., & Hoffman-Goetz, L. (2011). Predicting health literacy among English-as-a-second-language older Chinese immigrant women to Canada: Comprehension of colon cancer prevention information. *Journal of Cancer Education*, *26*(2), 326–332. doi: 10.1007/s13187-010-0162-2

Ubel, P. A., Smith, D. M., Zikmund-Fisher, B. J., Derry, H. A., McClure, J., Stark, A., … Fagerlin, A. (2010). Testing whether decision aids introduce cognitive biases: Results of a randomized trial. *Patient Education and Counseling*, *80*(2), 158–163. doi: 10.1016/j.pec.2009.10.021

Van Gerven, P. W. M., Paas, F., & Tabbers, H. K. (2006). Cognitive aging and computer-based instructional design: Where do we go from here? *Educational Psychology Review*, *18*, 141–157. doi: 10.1007/s10648-006-9005-4

Van Tulder, M., Furlan, A., Bombardier, C., & Bouter, L. (2003). Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine*, *28*, 1290–1299. doi: 10.1097/01.brs.0000065484.95996.af

Van Weert, J. C. M., Van Noort, G., Bol, N., Van Dijk, L., Tates, K., & Jansen, J. (2011). Tailored information for cancer patients on the Internet: Effects of visual cues and language complexity on information recall and satisfaction. *Patient Education and Counseling*, *84*(3), 368–378. doi: 10.1016/j.pec.2011.04.006

Volk, R. J., Jibaja-Weiss, M. L., Hawley, S. T., Kneuper, S., Spann, S. J., Miles, B. J., & Hyman, D. J. (2008). Entertainment education for prostate cancer screening: A randomized trial among primary care patients with low health literacy. *Patient Education and Counseling*, *73*(3), 482–489. doi: 10.1016/j.pec.2008.07.033

Walker, D., Adebajo, A., Heslop, P., Hill, J., Firth, J., Bishop, P., & Helliwell, P. S. (2007). Patient education in rheumatoid arthritis: The effectiveness of the ARC booklet and the mind map. *Rheumatology*, *46*(10), 1593–1596. doi: 10.1093/rheumatology/kem171

Watkins, I., & Xie, B. (2014). eHealth literacy interventions for older adults: A systematic review of the literature. *Journal of Medical Internet Research*, *16*(11), e225. doi: 10.2196/jmir.3318

Wilson, E. A. H., & Park, D. C. (2008). A case for clarity in the writing of health statements. *Patient Education and Counseling*, *72*(2), 330–335. doi: 10.1016/j.pec.2008.02.008

Wolf, M. S., Gazmarian, J. A., & Baker, D. W. (2005). Health literacy and functional health status among older adults. *Archives of Internal Medicine*, *165*(17), 1946–1952. doi: 10.1001/archinte.165.17.1946

Wolfson, N. E., & Kraiger, K. (2014). Cognitive aging and training: The role of instructional coherence and advance organizers. *Experimental Aging Research*, *40*(2), 164–186. doi: 10.1080/0361073x.2014.882206

Xie, B. (2011). Experimenting on the impact of learning methods and information presentation channels on older adults' e-health literacy. *Journal of the American Society for Information Science and Technology*, *62*(9), 1797–1807. doi: 10.1002/asi.2157510.1002/asi.21575

Zamarian, L., Benke, T., Buchler, M., Wenter, J., & Delazer, M. (2010). Information about medications may cause misunderstanding in older adults with cognitive impairment. *Journal of the Neurological Sciences*, *298*, 46–51. doi: 10.1016/j.jns.2010.08.061

Zamora, H., & Clingerman, E. M. (2011). Health literacy among older adults: A systematic literature review. *Journal of Gerontological Nursing*, *37*(10), 41–51.

Zikmund-Fisher, B. J., Fagerlin, A., Roberts, T. R., Derry, H. A., & Ubel, P. A. (2008). Alternate methods of framing information about medication side effects: Incremental risk versus total risk of occurrence. *Journal of Health Communication*, *13*(2), 107–124. doi: 10.1080/10810730701854011

Appendix A: Search Strategy for MEDLINE

MEDLINE (Via EBSCOHost), Date: 03-25-2013 (Exploratory Search)

No.	Strategy	Filters	Hits
S1	(MH "Health Literacy")	—	1189
S2	(MH "Reading")	—	16.104
S3	(MH "Educational Status")	—	37.353
S4	(TI (illiterac* OR illiterat* OR literate OR literac*) OR AB (illiterac* OR illiterat* OR literate OR literac*)) OR (S1 OR S3)	—	45.271
S5	(MH "Health Education") OR (MH "Patient Education as Topic")	—	116.258
S6	(MH "Consumer Health Information")	—	1479
S7	(MH "Patient Participation")	—	16.807
S8	(MH "Informed Consent+")	—	34.067
S9	(MH "Drug Prescriptions")	—	21.181
S10	(MH "Food Labeling")	—	2131
S11	(MH "Decision Making")	—	81.863

(Continued)

No.	Strategy	Filters	Hits
	OR (MH "Choice Behavior")		
S12	(MH "Pamphlets")	—	3000
S13	(MH "Comprehension")	—	7129
S14	(MH "Health Knowledge, Attitudes, Practice")	—	66.666
S15	(MH "Learning")	—	43.067
S16	TI (comprehen* OR readab* OR "ability to read" OR understand* OR processing OR knowledge OR learning OR performance OR accuracy OR error*) OR AB (comprehen* OR readab* OR "ability to read" OR understand* OR processing OR knowledge OR learning OR performance OR accuracy OR error*)	—	2.052.654
S17	"health advice*" OR "health message" OR "decision aid*" OR "health document*" OR "health text*" OR pamphlet OR booklet OR leaflet OR "patient information" OR "health information" OR "health related information"	—	33.743
S18	S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S17	—	283.573
S19	S13 OR S14 OR S15 OR S16	—	2.099.977
S20	S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S12 OR S17	—	212.257
S21	S20 AND S4 AND S19	—	2825
S22	S18 AND S4 AND S19	—	3101
S23	S20 AND S2 AND S19	—	379
S24	S22 OR S23	—	3270
S25	S22 OR S24	Age Related: All Child: 0–18 years	898
S26	S22 OR S24	Age Related: All Adult: 19 + years	2054
S27	S24 NOT (S25 NOT S26)	—	3034

The search strategy was adapted to suit the other databases. MH = Mesh Heading; TI = Title; AB = Abstract.

Appendix B: Acceptable Outcome Measures

We hereby provide a list of acceptable indicators of comprehensibility/comprehension. The list is based on indicators of comprehensibility and comprehension that are used in Database Begrijpelijke Taal NWO (Comprehensible Language and Effective Communication) and that appear in articles in this database that use measures other than the main ones listed. Note that this list is not exhaustive. It is possible that during our search we came across studies that used measures that could possibly have counted as indicators of comprehensibility/comprehension that are not listed here. Reviewers would have then discussed whether these indicators would have been acceptable.

- Attitudes about material (ease of reading, length of time needed to study material, self-reported comprehensibility/comprehension, self-reported confidence in learning, perceived utility of the material)
- Cloze test (in a cloze test, readers are asked to fill in every *n*th word that is left blank in the text)
- Correlation of experts' cognitive structures of material with subjects' cognitive structures of material
- Estimates of risk (self-reported estimates of risk after being presented with risk information)
- Expert assessment
- Expert opinions
- Finding information (speed of finding information, finding information)
- Free reproduction
- Guided reproduction
- Interpretation scores
- Interview with open-ended questions
- Pre- and posttest knowledge
- Multiple-choice questions for explicit information
- Multiple-choice questions for implicit information
- Open-ended questions for explicit information
- Open-ended questions for implicit information
- Questions for comprehension
- Reading (whether subjects do or do not read the passage)
- Reading time
- Recall of content (free or cued)
- Recall of text structure (free or cued)
- Recognition of the stimulus
- Recognition of related material
- Summary or paraphrase
- Task execution (immediate or delayed)
- Task performance (speed of performance)
- Task performance (accuracy of performance)
- Transfer of knowledge (using knowledge in a novel scenario)
- Word-naming task (speed of performance indicating comprehension)