Improving Patient Education Materials: A Practical Algorithm from Development to Validation

Patrick Betschart*  Sergej E. Staubli*  Valentin Zumstein  Christa Babst
Rafael Sauter  Hans-Peter Schmid  Dominik Abt

*Department of Urology; †Clinical Trials Unit, St. Gallen Cantonal Hospital, St. Gallen, Switzerland

Abstract

Objectives: To generate an algorithm for systematic development and validation of written patient information in accordance with well-established and validated psychometric and statistical methods that can be applied to different fields of medicine. Methods: A literature search was carried out in PubMed and Google Scholar. Methods were selected and combined to an algorithm. Feasibility and practicability is tested by the development of patient education materials on “ureteral stenting”. Results: The algorithm includes 4 study phases. After internal audit expert, readability of the first version is objectified using the Flesch Reading Ease formula. This draft is tested by a few patients performing semi-structured interviews using “The think aloud method” by Someren et al. Content validity is evaluated by a written survey by external consultants in accordance with Lawshe’s “Quantitative approach to content validity”. The final leaflet is developed at a consensus meeting and validated by patients based on the Consumer Information Rating Form. The new algorithm could be tested by the development of patient education materials on “ureteral stenting” as a test run. Conclusion: We developed an algorithm for systematic development and validation of written patient information in accordance with well-established, validated psychometric and statistical methods. This algorithm can be applied to arbitrary fields of medicine.
idated psychometric and statistical methods that can be applied to different fields of medicine. Morbidity associated with ureteral stenting represents a major concern in urology [21, 22]. It has been shown that stent-associated morbidity can be significantly reduced by high-quality patient education and that many patients are not satisfied with the information they obtained [3]. Therefore, “ureteral stenting” served as a topic for a proof-of-principle assessment of the developed algorithm.

Material and Methods

A comprehensive literature search for articles and methods published up to July 2018 was carried out in PubMed and Google Scholar using different combinations of the following search terms: “patient education material”, “patient information”, “validation”, “content”, “readability”, “comprehensibility”, “assessment”, “development”, “psychometric”, “statistical analysis”, “testing”, “leaflet”. We reviewed the publications found in this way, as well as the references cited in these publications.

Appropriate methods were selected and combined to an algorithm to guide through the process of initial text development, early assessment of content validity and comprehensibility, validation by patient survey and revisions that might be required after failure of validation. Feasibility and practicability of the developed algorithm was tested using the development of PEM on “ureteral stenting” as a proof-of-principle assessment. The study was approved by the local ethic committee and all participants provided written informed consent. Statistical analyses were performed in the R programming language (version 3.2.2) [23].

Results

Review

The literature search revealed several methods applicable for the development and assessment of PEMs. Methods used for the algorithm included the Flesch Reading Ease formula [24] (version adapted for German language by Amstad [25]), “The think aloud method” by Someren et al. [26], Lawshe’s “Quantitative approach
to content validity” [27], and the Consumer Information Rating Form (CIRF) [28]. These assessments were selected based on practicability, validity and their performance in previous studies and are described and discussed below.

Algorithm Development
The developed algorithm included 4 study phases and is illustrated in figure 1.

Phase 1: Initial Drafting of Text and Design The initial text and design are developed based on current literature available on the topic (e.g. reviews on indications, outcomes and adverse events). Particular attention should be paid to an easy comprehensibility of text and illustrations and an appealing design and layout. The initial version undergoes an internal audit by experts not involved into the initial drafting. Corrections are considered and consolidated at a first consensus meeting. Readability is then objectified using the Flesch Reading Ease formula [24]. A score between 60 and 69 points corresponding to a text with medium to rather easy understanding is defined as appropriate.

Phase 2: Initial Testing by Patients and Experts The first draft leaflet is tested by a few patients, who preferably have a history of the corresponding medical intervention or problem. Semi-structured interviews are conducted and recorded, using “The think aloud method” by Someren et al. [26].

In addition, content validity is evaluated by a written survey by external consultants in accordance with Lawshe’s “Quantitative approach to content validity” [27]. Each section of the PEM can be rated as “essential”, “useful but not essential” or “not necessary”, and the content validity ratio (CVR) is calculated according to the formula CVR = (ne – N/2) / (N/2) [ne: number of experts answering with “essential”; N: total number of experts interviewed]. With a possible range between +1 and -1, positive numbers indicate that at least 50% of all experts rate a chapter as “essential”, and the total CVR is calculated as the median of all ratios. For example, with 10 experts, a total CVR of 0.62 indicates validity, while lower values require content revision. Both patients and experts can give additional suggestions and free comments.

A final version for the validation process is drafted at a second consensus meeting. Re-assessment of readability as described above is performed if relevant changes were made to the text [24].

Phase 3: Validation The 17-item questionnaire CIRF is used for the final validation process, which is performed in patients with a history of the corresponding medical intervention or problem, sufficient language knowledge, and absence of cognitive impairment. The CIRF covers the topics comprehensibility (5 items), utility (6 items) and design quality (6 items). Items concerning comprehensibility and design quality are rated on a Likert scale from 1 (bad) to 5 (good) and items regarding utility are rated on a Likert scale from 1 to 4.

Phase 4: Revision and Repetition of Validation If the required validation threshold is not met, the respective item has to be adjusted and the validation process must be repeated after adjustment. If repetition of one or multiple items is triggered, the sample size has to be adjusted as the respective threshold for the p-value, corrected for multiple testing according to Bonferroni, depending on the number of items, which need reassessment (explained in the section “Algorithm Testing” and table. 1).

Algorithm Testing
The algorithm (fig. 1), we tested in a proof-of-principle assessment by creating and assessing a foldable 3-unit leaflet with a total of 6 A5 pages on ureteral stenting. The first draft was built in accordance to phase 1.

The title page contained the heading (“ureteral stent, patient information”), subtitle (“Information leaflet on frequently asked questions”) and a portrait of two neutral patients (male and female). The actual content was allocated to pages 2 to 5, including 3 photos (cystoscope,
stone in the ureter, unused ureteral stent) and 3 schematic illustrations of the urinary tract with and without an indwelling stent. The content was subdivided into the following sections: Why do you need a ureteral stent? All about the procedure (Procedure of ureteral stenting; After the intervention; What you should consider); Symptoms caused by the stent; Warning signs. Page 6 was used for contact details of whom to contact for further treatment and in case of problems, and included free space for contact details or a stamp.

The assessment of comprehensibility revealed an average sentence length of 10.11 words and an average number of syllables per word of 1.86. Thus, the Flesch Reading Ease formula [24] (adapted for German language by Amstad) [25] was calculated as follows: 180 \( - 10.11 \times (58.5 \times 1.85) = 61.67 \). This result corresponds to medium to rather easy readability, which was the target for the final version. The same category was reached after minor revisions in phase 2.

In general, the feedback from all 10 patients in phase 2 was very positive. However, some minor suggestions – usually words that were felt to be difficult to understand – were implemented, based on the patient interviews. According to a written survey by the external consultants, a high CVR (median 1, range 0.8–1) was found for all of the leaflet’s sections. Again, as a consequence of the experts’ free comments, some minor changes regarding the wording were implemented at a second consensus meeting.

For the validation phase, the second item of the CIRF subsection “utility” which addresses “contraindications” did not seem to be appropriate for ureteral stenting and was therefore deleted from the questionnaire. PEM was determined to be valid with respect to each item if more than 60% of the patients give an answer in the upper half of the Likert scale, i.e. 4 or 5 points for comprehensibility and design quality, and 3 or 4 points for utility. This hypothesis is examined by a one-sided binomial test for the

| CRF item | Description                                                                 | Number of answers in upper half of Likert scale | Proportions (%) | 95% CI         | P       |
|----------|-----------------------------------------------------------------------------|-------------------------------------------------|-----------------|----------------|---------|
| C1       | How easy would you say the information in the leaflet is to:                | 44                                             | 100.0           | 91.97–100.00   | < 0.0001 |
| C2       | Read                                                                        | 44                                             | 100.0           | 91.97–100.00   | < 0.0001 |
| C3       | Understand                                                                  | 39                                             | 88.6            | 76.02–95.05    | < 0.0001 |
| C4       | Remember                                                                    | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| C5       | Keep for the future                                                         | 40                                             | 90.9            | 78.84–96.41    | < 0.0001 |
| U1       | Indicate your opinion about how much information was provided and how useful you think this information would be: Benefits of the intervention | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| U2       | Contraindications of the intervention                                        | not applicable                                 | not applicable  | not applicable | not applicable |
| U3       | Providing contact information                                               | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| U4       | Precautions that need to be taken                                            | 43                                             | 93.2            | 81.77–97.65    | < 0.0001 |
| U5       | Adverse effects                                                             | 39                                             | 88.6            | 76.02–95.05    | < 0.0001 |
| D1       | Indwelling time of the stent                                                | 39                                             | 88.6            | 76.02–95.05    | < 0.0001 |
| D2       | Indicate your opinion about the following points concerning the design and structure quality of the leaflet | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| D3       | Organization/structure                                                       | 42                                             | 95.5            | 84.87–98.74    | < 0.0001 |
| D4       | Attractiveness/clearness                                                     | 44                                             | 100.0           | 91.97–100.00   | < 0.0001 |
| D5       | Print size of the text                                                       | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| D6       | Tone/configuration of the text                                               | 43                                             | 97.7            | 88.19–99.88    | < 0.0001 |
| D7       | Text spacing                                                                 | 43                                             | 93.2            | 81.77–97.65    | < 0.0001 |
| D8       | General design                                                               | 41                                             | 93.2            | 81.77–97.65    | < 0.0001 |

C = Comprehensibility; U = utility; D = design quality.
A 95% Wilson confidence interval (CI) and the p-value for the one-sided null hypothesis that the observed proportion is less or equal to 60% is provided.

Table 2. Proportions of patients providing answers in the upper half of the Likert scale for each CIRF item in phase 3

---

Improving Patient Education Materials

Curr Urol 2019;13:64–69

67
The think aloud
How well do you remember the content?

The algorithm ensures, that PEM that is thoroughly prepared. As con

If the required validation threshold – a p < 0.3125 – is not met, i.e. more than 8 patients provided an answer in the lower half of the Likert scale, the respective item was to be adjusted.

All of the 44 patients included in the study completed the CIRF and excellent validation results were found (table 2): with 44 responses (100%) in the upper half of the Likert scale, the best result was achieved for the items C1 (How readable is the text with regard to ease of reading?), C2 (How understandable is the text written?) and D3 (How easy is the font size to read?). The worst ratings with 39 answers (88.6%) in the upper half were the items C3 (How well do you remember the content?) and U5 (How do you find the information on the length of time the ureteral stent stays in place?).

For all 16 questionnaire items, more than 36 out of 44 patients (validation threshold) provided answers in the upper half of the Likert scale (median 97.7%, range 88.6–100%), and p-values for all items were < 0.0001, with respect to the one-sided null-hypothesis that the proportion with answers in the upper half of the Likert scale would be ≤ 60% (table 2).

As the p-values for all items were below the required validation threshold of 0.05/16, phase 4 adjustments and re-evaluation of the leaflet were not indicated. Number of patients and p-values that would have been required in case of failure of validation are provided in table 1.

Discussion

Comprehensibility of PEMs can only be assessed conclusively by patient surveys. Such validations are laborious and therefore, materials going into the validation process should already be thoroughly prepared. As content validity cannot be assessed by patients, this preparation should also include a review by specialists.

The developed algorithm ensures, that PEM that is used for validation already provides a high comprehensively, readability and content-related quality. It also demonstrates how validation can be performed and how to proceed in case of validation failures. Moreover, statistical concepts are provided for each step.

Our extensive literature review provided rather few methods for the development and validation of PEMs. Nevertheless, some of the methods used in the algorithm could be substituted by similar assessments. Though readability can be assessed by a variety of well-established methods (e.g. Flesch Reading Ease formula, Flesch-Kincaid Grade Level, SMOG Grade Level, Coleman-Liau Index, Gunning Fog Index), we recommend the use of the Flesch Reading Ease formula [24] as this is widely used for the assessment of technical manuals and medical content, easy to handle, and already integrated in the most popular word processing programs.

Though any kind of patient interview might serve as a rough indicator for comprehensibility, “The think aloud method” by Someren et al. [26] provides the advantage of assessing the patients first thoughts without being prone to restraint and avoidance of criticism by speaking aloud any words and ideas in ones mind while completing a task, for example reading a text.

Lawshe’s “Quantitative approach to content validity” [27] is a quite well-known method assessing the question of how far a group of experts agree on whether the knowledge measured by an indicator is “essential”, “useful but no essential” or “not necessary” for the measurement of the construct and there are few validated alternatives to this method.

For the final validation we used the CIRF [28], as this is a validated questionnaire which is specific for the assessment of written medical information and includes the whole spectrum of comprehensibility, utility and design of PEMs.

In summary, our selection of methods was based on an extensive literature research and all methods are all validated and generally accepted. Moreover, it is clearly more comprehensive than a survey for the evaluation of PEM leaflets published in 2003 [29].

While the algorithm might seem laborious and time consuming at a first glance, each of the steps can be performed quite quickly and the numbers of patients and experts can be easily adapted based on statistical considerations. In accordance, all of the steps performed in phase 1 and 2 during our pilot run could be performed with little efforts and provided important improvements of the PEM. As a consequence all of the CIRF items met the required validation threshold and no adaptions were necessary, which are likely to raise higher efforts than the simple tests performed in phase 1 and 2.
Improving Patient Education Materials

Conclusion

We developed an algorithm for systematic development and validation of written patient information in accordance with well-established and validated psychometric and statistical methods. This algorithm can be applied to arbitrary fields of medicine and proved to be easy to use and functional.

Acknowledgments

The authors thank Olympus Schweiz AG (Volketswil, Switzerland) for technical support and distribution of the leaflet, especially Petra Rufer, Esther Isenschmid and Ulrik Raether. In German speaking countries, interested physicians can obtain the leaflet on a non-commercial basis via their local Olympus representative. We also thank Alistair Reeves for text editing of the manuscript and the external consultants as well as the Clinical Trials Unit of Cantonal Hospital St. Gallen for their valuable input.

References

1. ComScore Reston, United States: Mobile health information audience jumps 125 percent in the past year. http://www.comscore.com/2012/01/mobile-health-information-audience-jumps-125-percent-in-the-past-year, cited July 2018.
2. Madden M, Fox S: Finding answers online in sickness and in health. Pew Internet and American Life Project 2006. http://www.pewinternet.org/2006/05/02/finding-answers-online-in-sickness-and-in-health/
3. Abt D, Warzinek E, Schmid HP, Haile SR, Engeler DS: Influence of patient education on morbidity caused by ureteral stents. Int J Urol 2015;22:679–683.
4. Coulter A: Evidence based patient information: important, so there needs to be a national strategy to ensure it. BMJ 1998;317:225–226.
5. Collodge A, Car J, Donnelly A, Majeed A: Health information for patients: time to look beyond patient information leaflets. J R Soc Med 2008;101:447–453.
6. Coulter A, Ellins J: Effectiveness of strategies for informing, educating, and involving patients. BMJ 2007;335:24–27.
7. Simpson M, Buckman R, Stewart M, Maguire P, Lipkin M, Novack D, Till J: Doctor-patient communication: the Toronto consensus statement. BMJ 1991;303:1385–1387.
8. Korsch BM, Negrete VF: Doctor-patient communication. Sci Am 1972;227:66–74.
9. Johansson K, Leino-Kilpi H, Salantera S, Lehtikunnas T, Ahonen P, Elomaa L, Salmela M: Need for change in patient education: a Finnish survey from the patient’s perspective. Patient Educ Couns 2003;51:239–245.
10. Ley P: Giving information to patients; in Eiser JR (ed): Psychology and Behavioural Medicine. Chichester, John, Wiley and Sons, 2008.
11. Guan Y, Maloney KA, Roter DL, Pollin TJ: Evaluation of the informational content, readability and comprehensibility of online health information on monogenic diabetes. J Genet Couns 2018;27:608–615.
12. Smith F, Carlsson E, Kokkinakis D, Forsberg M, Kodeda K, Sawatzky R, Friberg F, Ohlen J: Readability, suitability and comprehensibility in patient education materials for Swedish patients with colorectal cancer undergoing elective surgery: a mixed method design. Patient Educ Couns 2014;94:202–209.
13. Medlineplus: How to write easy-to-read health materials. Rockville Pike, United States. https://medlineplus.gov/etr.html.
14. Weiss BD: Health Literacy and Patient Safety: help patients understand. Manual for Clinicians (ed 2). Chicago, American Medical Association Foundation and American Medical Association, 2007.
15. Eloy JA, Li S, Kasabwala K, Agarwal N, Hunsberry DR, Baredes S, Setzcn M: Readability assessment of patient education materials on major otolaryngology association websites. Otolaryngol Head Neck Surg 2012;147:848–854.
16. Betschart P, Zumstein V, Bentivoglio M, Engeler D, Schmid HP, Abt D: Readability assessment of online patient education materials provided by the European Association of Urology. Int Urol Nephrol 2017;49:2111–2117.
17. Prabhu AV, Hansberry DR, Agarwal N, Clump DA, Heron DE: Radiation oncology and online patient education materials: deviating from NIH and AMA recommendations. Int J Radiat Oncol Biol Phys 2016;96:521–528.
18. Korsch BM, Linzer FM, Zeng-Treitler Q: Creating a gold standard for the readability measurement of health texts. AMIA Annu Symp Proc 2008;2008:353–357.
19. Michielutte R, Bahnson J, Dignan MB, Schroeder EM: The use of illustrations and narrative text style to improve readability of a health education brochure. J Cancer Educ 1992;7:251–260.
20. Houts PS, Doak CC, Doak LG, Loscalzo MJ: The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006;61:173–190.
21. Joshi HB, Okeke A, News N, Keeley FX Jr, Timoney AG: Characterization of urinary symptoms in patients with ureteral stents. Urology 2002;59:511–516.
22. Staubli SE, Mordasini L, Engeler DS, Sauter R, Schmid HP, Abt D: Economic aspects of morbidity caused by ureteral stents. Urol Int 2016;97:91–97.
23. R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015. https://www-r-project.org/.
24. Flesch R: A new readability yardstick. J Appl Psychol 1948;32:221–233.
25. Amstad T: Wie verständlich sind unsere Zeitungen (Dissertation)? Universität Zürich, 1978.
26. van Someren MW, Barnard YF, Sandberg JA: The think aloud method – a practical guide to modelling cognitive processes. London, Academic Press, 1994.
27. Lawshe CH: A quantitative approach to content validity. Pers Psychol 1975;28:563–575.
28. Koo MM, Krass I, Aslani P: Evaluation of written medicine information: validation of the Consumer Information Rating Form. Ann Pharmacother 2007;41:951–956.
29. Gustafsson J, Kalvmark S, Nilsson G, Nilsson JL: A method to evaluate patient information leaflets. Drug Inf J 2003;37:115–125.