Readability of patient information and consent documents in rheumatological studies

Hamnes, Bente; van Eijk-Hustings, Yvonne; Primdahl, Jette

Published in:
BMC Medical Ethics

DOI:
10.1186/s12910-016-0126-0

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license
CC BY

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Readability of patient information and consent documents in rheumatological studies

Bente Hamnes1*, Yvonne van Eijk-Hustings2 and Jette Primdahl3,4,5

Abstract

Background: Before participation in medical research an informed consent must be obtained. This study investigates whether the readability of patient information and consent documents (PICDs) corresponds to the average educational level of participants in rheumatological studies in the Netherlands, Denmark, and Norway.

Methods: 24 PICDs from studies were collected and readability was assessed independently using the Gunning’s Fog Index (FOG) and Simple Measure of Gobbledygook (SMOG) grading.

Results: The mean score for the FOG and SMOG grades were 14.2 (9.0–19.0) and 14.2 (12–17) respectively. The mean FOG and SMOG grades were 12.7 and 13.3 in the Dutch studies, 15.0 and 14.9 in the Danish studies, and 14.6 and 14.3 in the Norwegian studies, respectively. Out of the 2865 participants, more than 57 % had a lower educational level than the highest readability score calculated in the individual study.

Conclusions: As the readability level of the PICDs did not match the participants’ educational level, consent may not have been valid, as the participants may have had a limited understanding of what they agreed to participate in. There should be more focus on the readability of PICDs. National guidelines for how to write clear and unambiguous PICDs in simple and easily understandable language could increase the focus on the readability of PICD.

Keywords: Informed consent document, Readability, Gunning’s Fog Index, Simple Measure of Gobbledygook

Background

This study examines the readability of written patient information and consent documents (PICDs) used in rheumatology research. Ethical guidelines for medical research on human beings, which includes PICDs, were developed by the World Medical Association in Helsinki in 1964 [1]. The guidelines has been revised several times, the last time in 2013 [2]. The guidelines have also been implemented by National Ethic Committees [3–5].

Participants in medical research must give their informed consent before they can participate in a study. The consent must be free and informed, which means that the person should not experience any kind of pressure and should have the necessary information about the research which is to be conducted. This comprises knowledge about the purpose of the research project, duration, methods and procedures, types of expected results, and planned dissemination. Additionally, participants should have knowledge about the study’s potential benefits and risks, or negative consequences such as discomfort. Finally, participants should also receive information about confidentiality and data handling, in addition to knowledge about, that they can withdraw from the study at any time without any explanation and without influence on their clinical care [2]. The information should be written and easy to understand and carry as little risk as possible of misunderstanding. In medical research on humans, PICDs are key requirements, which must be approved by ethical committees [3–5]. In 2012, the European Commission described the literacy level in the population as important. Poor literacy is a hidden problem in most European societies, where one in five...
young people and adults lacks basic literacy skills [6]. Health literacy means the ability to read and understand basic health information, make appropriate health decisions, and take appropriate actions accordingly [7, 8]. People with poor health literacy will probably not understand a PICD if the text is too complex. In addition to the complexity of the text, participants’ reading ability is an important factor in ensuring their understanding of the written information. This must be taken into account when developing written materials like PICDs to ensure that patients are provided understandable information about the research study. This will allow them to make an informed decision about study participation [9].

Readability describes the number of years of schooling necessary to read and understand a text and can be measured by standardized instruments. Gunning’s Fog Index (FOG) [10] and the Simple Measure of Gobbledygook (SMOG) [11] are among the most commonly used instruments [12]. SMOG has a high correlation with FOG (0.95–0.99) [13, 14]. Scores from the two indexes reflect the average years of schooling a person needs to be able to read and understand a text. A SMOG or FOG score of 5 is equal to 5 years of schooling. In the health literature, the recommended readability level ranges from grades 5 to 9 [15–17].

Often PICDs have higher readability level than the study population or recommended readability level [15, 16, 18, 19]. In a study with 40 PICDs in anaesthesiology research, the readability score for the FOG was 11.9. This score was significantly higher than the average literacy level of the Australian and New Zealand populations, where 44–46 % people have a literacy level below the required minimum to meet the complex demands of everyday life [16]. In a Croatian study a readability score of 13.25 was found in informed consent documents in diagnostic and therapeutic procedures where 80 % of the population older than 15 years had less education [19]. Another study of patient information leaflets on physiotherapy showed that only 11 out of 33 leaflets (33 %) met the recommended readability level of grade 9 [15].

In addition to the readability level, the content of the text is essential for understanding. Studies on how participants understand medical consent documents showed four sources of uncertainty: language, information about risks and hazards, the nature of the procedure, and the documents’ composition and format [20]. This is congruent with the perspective of research ethical committee members which concluded that language, structure and format of the patient information sheet should be improved [21].

A review of interventions used to improve participants’ understanding of informed consent in research showed that one-to-one discussion and participant feedback with a study team member improved the participant’s understanding of the PICD. Other interventions such as multi-media and enhanced PICDs showed mixed levels of improvement in participant understanding [22]. Latest reviews from 2012–14 on improving participants understanding of PICDs, showed that enhanced PICDs and extended discussions were most effective [23]. For older participants interactive multimedia and written materials which were easy-to-read, increased the patients understanding of the consenting procedure [24]. For people with low literacy the most effective strategy was talking one-to-one with a study member, but this result is based on a single study only [25].

In the literature, it is recommended to conduct readability tests on patient materials, but this is not yet common practice [26]. During the European League Against Rheumatism Congress in 2010 a workshop regarding readability stressed the importance of the topic and it was recommended to be followed up. This study aims to document whether the readability of PICDs corresponds with reported education or schooling level of the participants enrolled in rheumatological studies.

**Methods**

**Materials**

We identified relevant clinical studies in the rheumatology field from each of the authors’ home countries (the Netherlands, Denmark, and Norway). Studies were selected in the researchers’ own languages to ensure that at least one with the native language reviewed all consent documents in each language.

An initial search was performed in PubMed in July 2014: #1: rheumat* OR arthritis OR arthro* OR psoriatic OR ankylosing spondylitis OR Bechterew OR lupus OR fibromyal* OR (inflammatory AND joint) OR chronic widespread pain #2: (Norway OR Norwegian; Holland OR Dutch; Denmark OR Danish) #1 AND #2. Additionally, we used filters to identify clinical trials that were performed within the last 10 years on humans; we searched for trials whose full texts were available. To include enough Danish rheumatological studies, we repeated the search for Danish studies in Cumulative Index to Nursing and Allied Health Literature (CINAHL).

Inclusion criteria were studies published during the past 10 years (2004–2014), study population ≥ 18 years of age, and a patient population with a specific rheumatic disease. Exclusion criteria were register-based studies, literature reviews, or other studies where no patient information or consent forms were used, studies from other countries than the Netherlands, Denmark, and Norway, and studies where no information about educational level (i.e. years of education, literacy level, and other proxies) was reported.

The titles and abstracts of the identified studies were screened according to the admission criteria. If a study
was not excluded based on the title or the abstract, the full text was retrieved and assessed according to the admission criteria. All authors screened the retrieved titles, abstracts, and full text from their own countries.

The first authors of the included studies were contacted through e-mail. We contacted the authors of the most recent studies first. They were informed about the aims of our study and were asked whether they would share the patient information and consent forms used in their study. If the first author did not respond within a week, we contacted the last author who was considered to have overall responsibility for the study. For one Danish and one Norwegian study, the authors were contacted for information about participants’ education level, which was not stated in the study, although the authors knew it was collected [27, 28].

As we expected each author could possibly reuse parts of the text from previous PICDs, we included only one study from each first author unless we knew that different people had been responsible for the development of the PICD. When data from the same study was reported in several papers by different first authors, only the author from the most recent paper was contacted. We aimed to be able to retrieve patient information and consent forms from 10 studies from each country and included studies that fulfilled the selection criteria consecutively.

Analysis
To assess the readability of the material, we applied the FOG [10] and SMOG [11] to each set of PICD. FOG grading was calculated based on a sample of approximately 100 words. The average number of words per sentence was calculated by dividing the number of words by the number of sentences in the sample. Words with three or more syllables were considered hard words. The number of hard words were counted. If polysyllabic words were repeated, only the first instance was counted. Three-syllable words made up of a two-syllable word with endings like -e, -er, -es, and -ing were omitted. The summed number of words per sentence and hard words were then multiplied with 0.4, and this gave the final score for each sample of approximately 100 words. We used the first 100 words and last 100 words of the text for scoring FOG 1 and FOG 2. SMOG grading was performed on 10 consecutive sentences at the beginning, 10 in the middle, and 10 near the end of the text. The number of words with three or more syllables was counted (including repetitions). The square root of the number of polysyllabic words (of the nearest perfect square) was calculated and the number ‘3’ was added to reach the final score [11]. In case the text did not contain 30 sentences for the SMOG grading, we assessed 10 sentences from the beginning of the text, the following 10 sentences, and finally 10 sentences counted backwards from the end of the text. Thus, some sentences were included twice.

In case abbreviations were defined, the abbreviated version was not counted as a hard word no matter the number of syllables. Headings and subheadings in the text were calculated as sentences only in case they consisted of at least two words.

Each PICD was graded independently by hand by two of the authors. Each author scored all the PICDs from her own country and half of the PICDs from each of the two other countries. Next, the grades were compared. In case of discrepancy between the calculated score, the third author scored the material as well, and a final grading was reached by comparison and discussion.

Next, we compared the participants’ educational or schooling level with the readability grade of the PICD for each single study. Since the age for starting school and the educational system varies from country to country, the years of schooling were reported based on each country’s system (Table 1). The proportion and number of participants with an educational or schooling level lower than the highest score for the FOG or the SMOG were reported.

Frequency distribution, means, and standard deviation (SD) were calculated for the FOG and SMOG grades in the PICDs.

Results
Through our search strategy, we retrieved 721 Dutch studies, 256 Danish studies and 192 Norwegian studies. After the screening, we included 24 studies, of which seven were Dutch, seven were Danish, and ten were Norwegian. Five of the included studies were qualitative, and 19 were quantitative. The number of participants in the studies ranged from 13 to 333. A total of 2972 participants were included in the 24 studies. In 23 of these studies, 1619 participants out of 2865 participants (57 %) had a lower educational or schooling level than the highest scores required for the FOG or the SMOG. The percentage of participants with lower educational or schooling level than the required FOG and the SMOG scores ranged from 5 % to 84 % in the included studies (Table 1). In one of the 24 studies, it was not possible to calculate the proportion of participants with a lower educational level than that suggested in the FOG/SMOG scores, because the results were presented as mean values [29] (Table 1).

Education or schooling levels presented in the included studies were often graded in two or three levels. The highest educational level was often described to be more than 12 or 13 years of schooling or as ‘university level’.
<table>
<thead>
<tr>
<th>Study no</th>
<th>Study design</th>
<th>No of participants (n)</th>
<th>Mean age and sex</th>
<th>Participants years of schooling or education level</th>
<th>FOG grade 1</th>
<th>FOG grade 2</th>
<th>SMOG grade</th>
<th>Participants with lower educational level than required to understand the PICD % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [47]</td>
<td>Evaluation of self-management programs</td>
<td>19</td>
<td>21 years (n = 9) and 22 years (n = 10), female 84 %</td>
<td>Low 1&lt;sup&gt;a&lt;/sup&gt; Middle 13&lt;sup&gt;b&lt;/sup&gt; High 5&lt;sup&gt;c&lt;/sup&gt;</td>
<td>11.0</td>
<td>9.4</td>
<td>13</td>
<td>5 % (1)</td>
</tr>
<tr>
<td>2 [48]</td>
<td>Cross-sectional study</td>
<td>333</td>
<td>47 years, female 100 %</td>
<td>Low 15&lt;sup&gt;d&lt;/sup&gt; Middle 256&lt;sup&gt;e&lt;/sup&gt; High 62&lt;sup&gt;f&lt;/sup&gt;</td>
<td>10.1</td>
<td>9.0</td>
<td>13</td>
<td>5 % (15)</td>
</tr>
<tr>
<td>3 [49]</td>
<td>Randomised controlled study</td>
<td>199</td>
<td>62 years, female 65 %</td>
<td>Low 28&lt;sup&gt;g&lt;/sup&gt; Middle 79&lt;sup&gt;h&lt;/sup&gt; High 91&lt;sup&gt;i&lt;/sup&gt; Missing 1</td>
<td>11.7</td>
<td>15.5</td>
<td>13</td>
<td>54 % (107)</td>
</tr>
<tr>
<td>4 [50]</td>
<td>Randomised controlled study</td>
<td>82</td>
<td>48 years, female 92 %</td>
<td>Low/intermediate 50 High 32</td>
<td>15.6</td>
<td>11.1</td>
<td>13</td>
<td>61 % (50)</td>
</tr>
<tr>
<td>5 [51]</td>
<td>Randomised controlled study</td>
<td>203</td>
<td>In five groups: 39 to 43 years, female 92–100 %</td>
<td>Low 108 Middle 68 High 27</td>
<td>12.3</td>
<td>17.6</td>
<td>14</td>
<td>87 % (176)</td>
</tr>
<tr>
<td>6 [52]</td>
<td>Qualitative focus group study</td>
<td>20</td>
<td>57 years, female 75 %</td>
<td>Low 3&lt;sup&gt;rd&lt;/sup&gt; Middle 12&lt;sup&gt;th&lt;/sup&gt; High 5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>13.8</td>
<td>11.8</td>
<td>12</td>
<td>75 % (15)</td>
</tr>
<tr>
<td>7 [53]</td>
<td>Randomised controlled study</td>
<td>158</td>
<td>47 years, female 94 %</td>
<td>Primary (mean 7 years) Secondary (mean 12 years) 128 Tertiary (mean 17 years) 22</td>
<td>15.6</td>
<td>13.9</td>
<td>15</td>
<td>86 % (136)</td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 [54]</td>
<td>Randomised crossover study</td>
<td>20</td>
<td>67 years, female 100 %</td>
<td>Elementary school 5 High school 0 &lt;4 years at University level 9 4-6 years at University level 2 &gt;6 years at University level 4</td>
<td>16.5</td>
<td>12.9</td>
<td>14</td>
<td>70 % (14)</td>
</tr>
<tr>
<td>2 [55]</td>
<td>A prospective study</td>
<td>315</td>
<td>55 years, female 77 %</td>
<td>None 67 Vocational 102 Higher 146</td>
<td>18.7</td>
<td>16.8</td>
<td>16</td>
<td>54 % (169)</td>
</tr>
<tr>
<td>3 [56]</td>
<td>A qualitative interview study</td>
<td>16</td>
<td>50 years, female 75 %</td>
<td>No formal education 2 Minimal education 7 Secondary education 4 University 3</td>
<td>13.0</td>
<td>11.0</td>
<td>13</td>
<td>69 % (9)</td>
</tr>
<tr>
<td>4 [57]</td>
<td>A qualitative interview study</td>
<td>13</td>
<td>44 years, male 100 %</td>
<td>Minimal education 1 Secondary education 9 University 3</td>
<td>11.5</td>
<td>17.0</td>
<td>15</td>
<td>77 % (10)</td>
</tr>
<tr>
<td>5 [58]</td>
<td>Randomised controlled study</td>
<td>287</td>
<td>63 years, female 87 %</td>
<td>&lt;11 years 232 ≥11 years 55</td>
<td>13.6</td>
<td>17.9</td>
<td>15</td>
<td>81 % (232)</td>
</tr>
</tbody>
</table>
Table 1  Gunning’s Fog Index (FOG) and Simple Measure of Gobbledygook (SMOG) grade levels of studies and participants’ educational levels (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Study Type</th>
<th>Sample Size</th>
<th>Participants</th>
<th>Education Level</th>
<th>FOG Grade</th>
<th>SMOG Grade</th>
<th>Graduates</th>
<th>Ungraduates</th>
<th>Graduates %</th>
<th>Ungraduates %</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Qualitative focus group study</td>
<td>32</td>
<td>58 years, female 59%</td>
<td>Unskilled 3 Vocational 18 Undergraduate 6 Graduate 5</td>
<td>15.2</td>
<td>13.8</td>
<td>14</td>
<td>84% (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>A qualitative interview study</td>
<td>31</td>
<td>38–89 years, female 85%</td>
<td>&lt;10 years 4 10–12 years 7 &gt; 12 years 20</td>
<td>17.8</td>
<td>14.4</td>
<td>17</td>
<td>36% (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>Randomised controlled study</td>
<td>57</td>
<td>Group 1: 68 and Group 2: 69 years, female 56%</td>
<td>≤12 years 24 &gt; 12 years 33</td>
<td>16.9</td>
<td>11.4</td>
<td>13</td>
<td>42% (24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Randomised controlled study</td>
<td>141</td>
<td>58 years, female 69%</td>
<td>≤12 years 88 &gt; 12 years 53</td>
<td>19.0</td>
<td>15.1</td>
<td>15</td>
<td>63% (88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Randomised controlled study</td>
<td>95</td>
<td>49 years, female 35%</td>
<td>≤12 years 58 &gt; 12 years 37</td>
<td>16.6</td>
<td>13.8</td>
<td>15</td>
<td>61% (58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Comparing treatment</td>
<td>153</td>
<td>Group 1: 57 years, female 89%. Group 2: 51 years, female 53%</td>
<td>≤12 years 87 &gt; 12 years 65</td>
<td>15.2</td>
<td>18.7</td>
<td>14</td>
<td>57% (87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Randomised controlled study</td>
<td>68</td>
<td>50 years, female 68%</td>
<td>Elementary school 13 High school 36 College/university 19</td>
<td>13.3</td>
<td>13.6</td>
<td>14</td>
<td>72% (49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Randomised controlled study</td>
<td>135</td>
<td>44 years, all female</td>
<td>≤10 years 21 11–13 years 49 &gt; 13 years 53 Unknown 12</td>
<td>14.9</td>
<td>9.8</td>
<td>14</td>
<td>52% (70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Follow up study</td>
<td>134</td>
<td>55 years, female 87%</td>
<td>9 years 36 ≤ 12 years 43 &gt; 12 years 55</td>
<td>12.8</td>
<td>13.3</td>
<td>14</td>
<td>59% (79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>A longitudinal study</td>
<td>281</td>
<td>46 years, female 58%</td>
<td>≤9 years 38 10–12 years 118 &gt; 12 years 125</td>
<td>12.8</td>
<td>13.7</td>
<td>15</td>
<td>56% (156)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Randomised controlled study</td>
<td>107</td>
<td>M-group 48 years, N-group S1ys, female 43%</td>
<td>Mean years of education 13 years (M-group) 12 years (N-group)</td>
<td>15.0</td>
<td>18.7</td>
<td>15</td>
<td>Cannot be calculated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Randomised controlled study</td>
<td>73</td>
<td>54 years, female 79%</td>
<td>≤12 years 36 &gt; 12 years 37</td>
<td>12.9</td>
<td>13.6</td>
<td>14</td>
<td>49% (36)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Participants in 23 studies 28653

Explanation of years of schooling or education level (a–f)

aVocational training
bAdvanced vocational training
cCollege/University training
dPrimary school/low vocational secondary education
eIntermediate general secondary education/vocational education
fHigher general secondary education, higher vocational education, preuniversity or university level
gWithout the Norwegian study number 9 which cannot be calculated
All PICDs were written in standard templates without illustrations, figures or other initiatives that could improve the readability. During the analysis, there were minor discrepancies, which were discussed. The reasons were often words that were overlooked, or that the pronunciations were unknown for one of the authors and thus the number of syllables could be interpreted differently.

The total mean score for the FOG and SMOG grades were 14.2 (9.0–19.0) and 14.2 (12–17) respectively. The mean scores for the FOG and SMOG were 12.7 and 13.3 in the Dutch studies, 15.0 and 14.9 in the Danish studies, and 14.6 and 14.3 for the Norwegian studies, respectively (Table 2).

Discussion

This study investigated whether the readability of the PICDs corresponds with the reported education or schooling level of the participants involved in rheumatological studies in the Netherlands, Denmark, and Norway.

The results show that in all 24 studies, the PICDs had a higher readability level than recommended for health literature in general [15–17], and that in all studies there were participants with a lower educational or schooling level than the required readability level in order to be able to read and understand the PICD used in the study. This corresponds with findings from a new study of 522 information sheets for research studies in the UK [30].

In total 57 % of the participants had less education or schooling than required to read and understand the PICD based on the readability scores for the FOG and SMOG. When participants have a lower level of education than the readability of the PICD, it can lead to a ‘readability gap’, meaning that participants will not fully understand the text in the PICD or what they are agreeing to [31]. This may cause participants to refuse to participate in the study or imply that they are participating in the study without giving valid consent. Participants’ understanding of the text can be improved by oral information or discussion with a study team member about the content in the PICD [22]. It is also possible to increase PICDs’ readability by using short words and phrases, headings, plenty of space, bulleted lists, illustrations, and large type [20, 32, 33].

An individual’s health literacy depends on more than education; for example, it can depend on how familiar the reader is to the health care system and the presented information [34]. When readers possess such knowledge, they may understand the text in PCIDs even when their educational level is lower than the readability level of the PICD in question. Conversely, when readers lack this knowledge, they may fail to understand the text even when they their educational level corresponds to the readability level of the PICD.

The high readability level of the PICDs in this study can possibly be explained by the regulatory requirements in national guidelines, which give more attention to the content of the PICDs than their readability. In Denmark and Norway, there are no stated requirements or focus on the level of readability of the PICDs. In the Netherlands, the ethical committees recommend to write the material on a schooling level of lower secondary school, which is approximately at an age of 11–12 years. The literature has shown that the readability level of PICDs in USA, Australia, South Africa, India and several European countries does not fit requirements so it is plausible to suggest that it is a common and worldwide problem [16, 19, 21, 35–38]. Our findings and how we have conducted our study can therefore be relevant for health professionals working with PICDs.

For all the included studies, we received the PICD from the first author, and several authors commented that the current study was an important piece of work that they would gladly support. In this study, it was difficult to find 10 Dutch and 10 Danish studies in which education or years of schooling was reported as part of the participants’ demographic data. This indicate a lack of focus on the participants’ educational of schooling level, although social inequality is an important topic.

Patients’ educational level may affect how participants manage their disease and treatment [39]. Educational level can also influence the risk of disease [40], disease severity [41], psychological distress [42], health [43], involvement in healthcare [44], and mortality [45]. Therefore, it may be advantageous for studies to report the educational or schooling level of their participants.

This study’s strength is that it examines PICDs from studies in three different European countries. It enables us to obtain knowledge on the readability of PICDs in a larger context. Using two different tests for readability also strengthens the study. Average values for FOG and SMOG in the PICDs are quite similar in our study, but

<table>
<thead>
<tr>
<th></th>
<th>Number of values</th>
<th>Mean (min-max)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOG-NL</td>
<td>14</td>
<td>12.7 (9.0–17.6)</td>
<td>2.6</td>
</tr>
<tr>
<td>FOG-DK</td>
<td>14</td>
<td>15.0 (11.0–18.7)</td>
<td>2.5</td>
</tr>
<tr>
<td>FOG-NO</td>
<td>20</td>
<td>14.5 (9.8–19.0)</td>
<td>2.4</td>
</tr>
<tr>
<td>SMOG-NL</td>
<td>7</td>
<td>13.3 (12.0–15.0)</td>
<td>1.0</td>
</tr>
<tr>
<td>SMOG-DK</td>
<td>7</td>
<td>14.9 (13.0–17.0)</td>
<td>1.4</td>
</tr>
<tr>
<td>SMOG-NO</td>
<td>10</td>
<td>14.3 (13.0–15.0)</td>
<td>0.7</td>
</tr>
<tr>
<td>FOG-All</td>
<td>48</td>
<td>14.2 (9.0–19.0)</td>
<td>2.6</td>
</tr>
<tr>
<td>SMOG-All</td>
<td>24</td>
<td>14.2 (12.0–17.0)</td>
<td>1.1</td>
</tr>
</tbody>
</table>
within individual PICDs, there is a wide variation in scores. This shows that the complexity of PICD texts varies and that it is difficult to obtain a single readability measure of the PICDs. Readability tests like FOG and SMOG have been criticized for focusing on the document alone, and for not capturing medical terms and difficult words [12]. A weakness of a readability test is that it does not take into account the context in which the document is used [12].

For assessment of PICDs and other written patient information, there are several online tools which involve FOG, SMOG, and other readability tests. These electronic tools make it possible to quickly obtain the value of a text's readability. Online tools are not exact and provide somewhat different results than tests scored by hand. One such tool calculates a total readability score based on several tests (http://www.readabilityformulas.com). Another tool shows the readability for each test and provides a list of which sentences should be rewritten to improve readability (http://www.online-utility.org/english/readability_-_test_and_improve.jsp). This tool is used in an example of PICD where the same text is written with high and low readability, see Additional file 1.

In our study, several of the PICDs were only available on paper. We also wanted to calculate the readability level as exactly as possible. Therefore, all PICDs were scored by hand.

Study limitation
The studies included in this study use different ways of describing the participants' level of education. Some use the number of years of schooling, grouped or as average. Other studies use low-medium-high educational classifications, and some studies describe the professions or occupations of the participants. This makes it difficult to compare results. In most studies, there will be more participants who have lower educational levels than PICDs' readability level (Table 1), but the way participants' education is presented in the studies makes it impossible to perform a more accurate calculation. For example, when the number of participants in a study is presented with an education level of more than 12 years and the highest readability score is 19.0, then only a few in this group will have a schooling level equivalent to 19.0 years. A standardized way of presenting participants' educational level is needed.

The FOG and the SMOG were developed to score texts written in English, which together with Dutch, Danish, and Norwegian, belong to the Germanic language family [46]. All four languages have common language characteristics. This is why we assumed we could use the FOG and SMOG procedure in all three languages even though they are not yet validated in the three languages. The endings which are considered to be easy to understand varies between languages. We cannot ignore the fact that the scores may be influenced by the languages.

Conclusion
More than 57% of the participants in the included studies had a lower educational or schooling level than the readability level of the PICDs. This raises concerns about the validity of the participants' consent. It would be relevant to develop national recommendations for the readability of PICDs. In addition, the involvement of patient research partners may help ensure a highly necessary focus on participants' reading level. Further research is needed in other medical fields and in other countries.

Practice implications
Our findings show that there should be more focus on the formulation and form of PICDs. The development of guidelines for writing PICDs would ensure that focus is placed on the formulation of a text that is clear and understandable. The use of such a text along with oral information will enable participants to understand a study's contents and consequences and give valid consent. This will also ensure the best possible recruitment in the studies.

Additional file

Additional file 1: PICD with high and low readability (invented). (DOCX 15 kb)

Abbreviations
DK, Denmark; FOG, Gunning's fog index; NL, Netherlands; NO, Norway; PICD, Patient information and consent document; SMOG, Simple measure of gobbledygook.

Acknowledgements
We thank patient research partner Carole Lindsell for her valuable comments. We also thank all the following authors for sharing their PICDs with us: Judy Ammerlaan, Ecolle R. Bossera, Daniell Bossen, Nadia Garnefski Sabia van Koulil, Henrik Rindel Gudbergsen, Louise Linde, Katrine Bjerre Loeppenthin, Mette Madsen, Tine Mechlenborg Kristiansen, Elise Bromann Bukhave, Vigdis Bruun-Olsen, Kjersti Grønning, Ingvild Kjeken, Mari Klokkerud, Hege Svean Koksav, Ofir Birna Kristjánsdóttir, Ida Lachting, Maria Dahl Mjaaavaiten, Yndis Annette Staalesen Strumse, and Heidi Andersen Zangi.

Funding
We want to thank the Norwegian Nurses in Rheumatology and Medac Pharma, Norway for supporting our project meeting in Copenhagen. We also want to thank the Hospital for Rheumatic Diseases, Lillehammer, Norway for supporting our project.

Availability of data and materials
The data used to support the conclusions of this article is publicly available as cited in the manuscript. For two studies, the first author was contacted for information about participants’ education level. This information is available from the first authors of the two articles.
Author's contributions

BH designed the study. JP performed the literature search. All authors obtained informed consent documents and performed the analyses. BH, JP and YEH drafted the manuscript and all authors read and approved the final manuscript.

Competing interest

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

This study used the results of previously published studies. The first author of the manuscripts. Consent for publication

The authors declare that they have no competing interests.

References

3. Det etiske råd [http://www.etiskraad.dk/]
4. The National Committee for Medical and Health Research Ethics (NEM) [https://www.etikkom.no/en/]
5. Central Committee on Research Involving Human Subjects [http://www.ccmo.nl/]
6. Vassilou A: For 1 in 3 Europeans, the world is hard to read. In: European Commission High Level Group of experts on Literacy, Brussels/Nicolaia: 2012.
43. Brekke M, Hjordal P, Kvien TK. Involvement and satisfaction: a Norwegian study of health care among 1,024 patients with rheumatoid arthritis and