Sense and readability: participant information sheets for research studies

Liam Ennis and Til Wykes

**Background**
Informed consent in research is partly achieved through the use of information sheets. There is a perception however that these information sheets are long and complex. The recommended reading level for patient information is grade 6, or 11–12 years old.

**Aims**
To investigate whether the readability of participant information sheets has changed over time, whether particular study characteristics are related to poorer readability and whether readability and other study characteristics are related to successful study recruitment.

**Method**
We obtained 522 information sheets from the UK National Institute for Health Research Clinical Research Network: Mental Health portfolio database and study principal investigators. Readability was assessed with the Flesch reading index and the Grade level test.

**Results**
Information sheets increased in length over the study period. The mean grade level across all information sheets was 9.8, or 15–16 years old. A high level of patient involvement was associated with more recruitment success and studies involving pharmaceutical or device interventions were the least successful. The complexity of information sheets had little bearing on successful recruitment.

**Conclusions**
Information sheets are far more complex than the recommended reading level of grade 6 for patient information. The disparity may be exacerbated by an increasing focus on legal content. Researchers would benefit from clear guidance from ethics committees on writing succinctly and accessibly and how to balance the competing legal issues with the ability of participants to understand what a study entails.

**Declaration of interest**
None.

**Copyright and usage**
© The Royal College of Psychiatrists 2015. This is an open access article distributed under the terms of the Creative Commons Non-Commercial, No Derivatives (CC BY-NC-ND) licence.
100 words, the next subheading was removed and counting continued into the next section. Sections excluded lists or contact details.

Study characteristics
We also recorded:
(a) Clinical study group (CSG): the clinical area to which a study belonged. These are based on the strategic analysis of UK mental health research funding categories.16
(b) Level of patient involvement in the study: based on categories reported in Ennis & Wykes.14 These were consultation only, researcher initiated collaboration and user controlled/user initiated collaboration.
(c) Study complexity: on a 1–17 scale, with higher scores indicating greater complexity, calculated on factors such as number and frequency of follow-ups, number of study sites and involvement of patients who lack capacity.
(d) Intervention type: whether the study included an intervention, and if so, what that was.
(e) Recompense value: how much payment was received for participation.
(f) Whether a study had or was on target to recruit to time and target. This was a binary variable, with >90% indicating successful recruitment as this usually enables a valid test of the study hypothesis.

Statistical analysis
To assess whether information sheets had changed over 10 years, we calculated Spearman’s Rho between time and the readability measures.

To identify differences in readability measures between CSGs, we used one-way analysis of variance (ANOVA). The Games–Howell procedure was used to test differences in means in the presence of heteroscedasticity.

ANOVA was also used to compare readability measures between studies with different interventions and to investigate whether it was affected by different levels of patient involvement. The Games–Howell procedure was used for post hoc comparisons where necessary. Mean Flesch index scores for the most complex section were compared using Hochberg’s GT2.

Binary logistic regression identified associations between various predictors and successful recruitment. All putative study characteristic predictors were entered simultaneously (those identified by Ennis & Wykes,14 participant payment and type of intervention), along with the Flesch index score for the whole information sheet. Other readability measures were highly correlated so we chose the Flesch index because it has greater precision than the reading grade score.

Sample size
We used G*Power 3.1.9 for sample size calculations.
(a) Time x readability correlations: with a sample size of 472, we had 90% power to detect a correlation of 0.1, one-sided ($\alpha = 0.05$).
(b) Readability differences between CSGs: with a sample size of 37 per group, we had 90% power to detect an effect size of 0.3, two-sided ($\alpha = 0.05$).
(c) Readability differences between intervention types: with a sample size of 20 per group, we had 90% power to detect an effect size of 0.4, two-sided ($\alpha = 0.05$).
(d) Readability differences between levels of patient involvement: with a sample size of 120 per group, we had 90% power to detect an effect size of 0.2, two-sided ($\alpha = 0.05$).
(e) Associations with successful recruitment: with a sample size of 313, we had 90% power to detect an odds ratio of 1.5, two-sided ($\alpha = 0.05$).

Results
A total of 278 of 522 (53%) studies were observational. A total of 254 studies (48.7%) offered some recompense for participation. Studies that involved psychotic disorders or addictions were more likely to provide payment than studies in other CSGs ($\chi^2(4) = 47.60, P < 0.001$) (Table 1).

The mean number of words in an information sheet was 1527, but length varied widely (range 161–5407). The mean Flesch index score was 59 (range 29.2–92.0) for whole information sheets, grade level 10 (range 3–14) or 15–16 years old. For the most complex section, the mean Flesch index score was 42 (range 1.8–84.2), grade level 13 (range 5–27) or 18–19 years old.

### Have information sheets become more complex over time?
Information sheets significantly increased in length over 10 years ($r = 0.18, P < 0.001$). The increase is from an average of 1333 words in 2003 to 1714 words in 2013. This relationship was not explained by an increasing number of interventional studies over time. Despite increases in length, there was no change over time in Flesch reading index or grade level.

### Table 1 Study characteristics

<table>
<thead>
<tr>
<th>Intervention type, a n (%)</th>
<th>Psychological</th>
<th>157 (30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacological/device</td>
<td>41 (8)</td>
<td></td>
</tr>
<tr>
<td>Service</td>
<td>26 (5)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>20 (4)</td>
<td></td>
</tr>
<tr>
<td>Observational</td>
<td>278 (53)</td>
<td></td>
</tr>
<tr>
<td>Study complexity, b mean (s.d.)</td>
<td>9.59 (4)</td>
<td></td>
</tr>
<tr>
<td>Clinical study group, c n (%)</td>
<td>130 (25)</td>
<td></td>
</tr>
<tr>
<td>Psychotic disorders</td>
<td>179 (34)</td>
<td></td>
</tr>
<tr>
<td>Mood and personality disorders</td>
<td>80 (15)</td>
<td></td>
</tr>
<tr>
<td>Services research</td>
<td>76 (15)</td>
<td></td>
</tr>
<tr>
<td>Dementias and intellectual disability</td>
<td>37 (7)</td>
<td></td>
</tr>
<tr>
<td>Addictions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient involvement, d n (%)</td>
<td>200 (38)</td>
<td></td>
</tr>
<tr>
<td>Consultation</td>
<td>199 (38)</td>
<td></td>
</tr>
<tr>
<td>Researcher initiated collaboration</td>
<td>120 (23)</td>
<td></td>
</tr>
<tr>
<td>Jointly/patient initiated collaboration/patient control study</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recompense value, e £</th>
<th>Median</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median for those providing compensation</td>
<td>30.00</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>5–300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Readability statistics, f mean (s.d.)</th>
<th>Number of words</th>
<th>1527 (821)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flesch index: whole sheet</td>
<td>58.79 (9)</td>
<td></td>
</tr>
<tr>
<td>Grade level: whole sheet</td>
<td>9.84 (2)</td>
<td></td>
</tr>
<tr>
<td>Flesch index: most complex section</td>
<td>41.50 (13)</td>
<td></td>
</tr>
<tr>
<td>Grade level: most complex section</td>
<td>13.33 (3)</td>
<td></td>
</tr>
</tbody>
</table>

a. n = 522.
b. n = 520.
c. n = 522.
d. n = 519.
Do study characteristics affect readability?

Information sheets from different CSGs varied significantly in length as measured by number of words ($F_{(4, 497)} = 9.80$, $P < 0.001$), the Flesch index for the most complex section ($F_{(4, 479)} = 5.42$, $P < 0.001$) and the whole information sheet ($F_{(4, 497)} = 8.10$, $P < 0.001$). The mean scores for these measures as a function of CSG are presented in Table 2.

Information sheets for studies involving dementias and intellectual disabilities were the shortest and easiest to read of any CSG.

Information sheets for studies including different interventions varied in length ($F_{(4, 517)} = 58.00$, $P < 0.001$) and in Flesch index scores for the most complex section ($F_{(4, 517)} = 3.12$, $P = 0.015$). Studies including a pharmacological or device intervention were the longest and most complex. Unsurprisingly, information sheets for studies containing no intervention (i.e. observational studies) were the shortest. There was no significant difference on the Flesch index scores for the whole information sheet. The mean scores for all of these measures are presented in Table 2.

The length of information sheets was significantly different between studies which included different levels of patient involvement ($F_{(2, 516)} = 4.71$, $P = 0.009$). This was an inverse linear relationship; as patient involvement increased, length decreased by about 150 words. There was no significant difference in the other readability measures. Means are given in Table 2.

Does readability predict recruitment success?

The equation used to identify associations with successful recruitment was very close to statistical significance (deviance $\chi^2 (14) = 23.19$, $P = 0.057$). The model (Table 3) produced a Nagelkerke pseudo-$R^2$ of 0.064.

The model shows that studies which involved a pharmacological or device intervention were less likely to recruit to time and target than other types of study. Studies that involved patients at the highest level were more likely to achieve successful recruitment, but readability did not contribute to recruitment success.

Discussion

We have analysed a study sample nearly double the size of the next largest study (284 studies).17,18 We also covered the longest period, 10 years. This has allowed us to reveal some important information on the state of participant information sheets for today’s research studies.

Have information sheets changed over time?

Our data demonstrate that information sheets have grown longer over time. This may be a result of the increasing focus on patient safety – what some might term risk aversion – over the period of study.19,20 This is noteworthy because previous research found that more detailed information sheets are less well understood than briefer versions.11–13 Longer information sheets are also less likely to be read.2,4 Taken with these findings, our results imply that participants’ understanding may have actually decreased over time. This is a hypothesis that needs testing, as it has implications for ethics committee advice. On the positive side, information sheets do not appear to have become more complex over time.

Is readability affected by study characteristics?

Information sheets for studies investigating dementias and intellectual disabilities were easier to read than studies in some other areas. This is unsurprising, since many studies belonging to this CSG used simplified language supported with lots of pictures. However, even with this adjustment, information sheets in this CSG only crept into the ‘standard’ range of reading difficulty (Flesch index score 61–70),15 requiring an estimated reading grade of 9 or 13–14 years old. Information sheets for all other CSGs were firmly in the ‘fairly difficult’ category (Flesch index score 51–60). In addition, the mean score for the most complex section in every CSG fell into the ‘difficult’ category (Flesch index score 31–50).

Our data also show that the information sheets for observational studies were significantly shorter than some other types. This is unsurprising since observational studies need not describe interventions, the process of group allocation or masking procedures. Perhaps more interesting is the average length of information sheets for studies including a pharmacological or device intervention – a staggering 3000 words. Pharmacological and device intervention studies also performed poorly with regard to the most complex section, with the average score falling into the ‘very difficult’ category (0–30). Texts scoring in this range are comparable to scientific writings.15 This is of particular importance since more complex information sheets tend to accompany studies which carry the most risk.21

Table 2: Mean readability measures for different clinical study groups (CSG), intervention types and levels of patient involvement(a)

<table>
<thead>
<tr>
<th>CSG</th>
<th>Words, n</th>
<th>Flesch reading index</th>
<th>Reading grade level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Whole sheet</td>
<td>Most complex section</td>
</tr>
<tr>
<td>Dementia and intellectual disability</td>
<td>1297 (1155–1440)</td>
<td>63.59 (60.69–66.50)</td>
<td>47.67 (43.75–51.06)</td>
</tr>
<tr>
<td>Services research</td>
<td>1176 (1079–1273)</td>
<td>56.65 (55.09–58.21)</td>
<td>41.78 (39.38–44.17)</td>
</tr>
<tr>
<td>Addictions</td>
<td>2002 (1547–2457)</td>
<td>57.65 (55.57–59.73)</td>
<td>40.80 (38.61–44.79)</td>
</tr>
<tr>
<td></td>
<td>1267 (1200–1333)</td>
<td>58.48 (57.49–59.47)</td>
<td>41.86 (40.37–43.40)</td>
</tr>
<tr>
<td>Psychological</td>
<td>1547 (1504–1645)</td>
<td>59.75 (58.31–61.19)</td>
<td>42.88 (40.58–45.18)</td>
</tr>
<tr>
<td>Service</td>
<td>1410 (1175–1645)</td>
<td>59.80 (56.89–62.70)</td>
<td>41.94 (37.29–46.59)</td>
</tr>
<tr>
<td>Other</td>
<td>2318 (1745–2892)</td>
<td>59.54 (55.02–64.05)</td>
<td>37.52 (30.20–44.84)</td>
</tr>
<tr>
<td></td>
<td>1643 (1510–1775)</td>
<td>59.05 (57.79–60.30)</td>
<td>41.62 (39.77–43.48)</td>
</tr>
<tr>
<td>Consultation</td>
<td>1509 (1400–1617)</td>
<td>58.73 (57.46–59.99)</td>
<td>41.67 (39.76–43.57)</td>
</tr>
<tr>
<td>Researcher initiated collaboration</td>
<td>1355 (1243–1467)</td>
<td>58.33 (57.04–59.65)</td>
<td>41.09 (38.70–43.49)</td>
</tr>
</tbody>
</table>

**Table 2** Mean readability measures for different clinical study groups (CSG), intervention types and levels of patient involvement(a)

---

(a) 95% confidence intervals are shown in parentheses.
Higher levels of patient involvement seem to facilitate briefer information sheets. This could be a product of patients reviewing information sheets and commenting on sections which could be shortened. It is interesting to note, however, that patient involvement did not seem to mitigate against complex information sheets overall nor against very dense sections of text.

Collapsing the groups used in this study, the average whole-sheet Flesch index score was 59, corresponding to US grade 10, or a reading age of 15–16 years old. This is higher than the UK national reading age of US grade 8, or 13–14 years old.22 It is also markedly higher than the recommended reading age for patient information texts (US grade 6, or 10–11 years). About 89% of the information sheets in our sample were at or above the national reading age of US grade 8, or 13–14 years old.22 It is also higher than the UK national reading age of 15–16 years old. This is higher than the reading age for patient information sheets and commenting on sections which could be shortened. It is interesting to note, however, that patient involvement did not seem to mitigate against complex information sheets overall nor against very dense sections of text.

The most complex sections of information sheets were very demanding. On average they scored 42 on the Flesch index, which corresponds to a reading age of 18–19. One information sheet we analysed scored 73 (fairly easy) for the whole sheet but a dismal 12 (very difficult) for the most complex section, showing how important it is to consider the complex section in any analysis.

Despite the poor performance of the information sheets we analysed, the results were better than those reported by others. A study of mental health research recorded a mean Flesch index score of 48, or grade 12.4 Similarly, oncology studies scored 45,4 anaesthesia research a reading grade 12,10 and one cross-discipline French study actually reported a median Flesch index score of 24 (very difficult).4 We have examined the methods of these studies and they are comparable to our own. The different results might therefore be attributed to variation in the composition of study samples. It seems likely, for example, that a sample of anaesthesia studies would contain more clinical trials than our own sample.

There are a number of techniques to reduce the reading age of information sheets which are obvious – using shorter words, sentences and paragraphs and replacing complex medical and research terms with simple words.25 A similar glossary could be produced extending – and in some cases reducing – overall length. In one case, we were able to increase the Flesch index score from 18.9 (very difficult) to 73.6 (fairly easy).

### Does readability have implications for recruitment?

Our analysis showed that the complexity of information sheets had little bearing on successful recruitment. The odds ratio of one indicates that even very large changes in Flesch index score are unlikely to affect the chances of recruitment success. Our results therefore suggest that information sheet complexity, at least as measured by the Flesch index, remains an ethical problem rather than an ingredient for study success.

We found that studies including a pharmacological or device intervention were less likely to have reached their target than observational studies even in this non-commercial funded dataset. Overall, however, this finding will not come as a surprise to many readers: difficulty in recruiting to drug trials is an enduring issue.26–29 Qualitative studies have shown that some people dislike the idea of ‘being a guinea pig’,30 the rigidity of treatment regimes and the prospect of side-effects.31

### Limitations

There are two limitations to this study. The first is that the study did not investigate comprehension by real people but instead used

---

**Table 3 Logistic regression showing associations with successful recruitment**

<table>
<thead>
<tr>
<th>Clinical study group</th>
<th>Beta</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia and intellectual disability</td>
<td>−0.04</td>
<td>0.96</td>
<td>0.52–1.76</td>
</tr>
<tr>
<td>Mood and personality disorders</td>
<td>−0.37</td>
<td>0.69</td>
<td>0.34–1.42</td>
</tr>
<tr>
<td>Services research</td>
<td>−0.10</td>
<td>0.90</td>
<td>0.48–1.70</td>
</tr>
<tr>
<td>Psychotic disorders</td>
<td>0.44</td>
<td>1.56</td>
<td>0.63–3.83</td>
</tr>
<tr>
<td>Addictions</td>
<td>−0.02</td>
<td>0.99</td>
<td>0.94–1.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complexity</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient involvement</td>
<td>0.24</td>
<td>1.28</td>
<td>0.83–1.97</td>
</tr>
<tr>
<td>Consultation only</td>
<td>−0.51</td>
<td>1.66</td>
<td>1.00–2.77*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intervention type</th>
<th>Beta</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational</td>
<td>−0.19</td>
<td>0.83</td>
<td>0.51–1.34</td>
</tr>
<tr>
<td>Psychological</td>
<td>−1.44</td>
<td>0.24</td>
<td>0.10–0.54**</td>
</tr>
<tr>
<td>Pharmacological/device</td>
<td>0.12</td>
<td>1.13</td>
<td>0.45–2.84</td>
</tr>
<tr>
<td>Service</td>
<td>−0.50</td>
<td>0.61</td>
<td>0.22–1.70</td>
</tr>
<tr>
<td>Other</td>
<td>0.00</td>
<td>1.00</td>
<td>0.98–1.02</td>
</tr>
</tbody>
</table>

| Flesch index (whole sheet)                         | 0.05    | 1.01       | 1.00–1.01|
| Recompense value                                   | 0.00    | 1.00       | 1.00–1.00|

| Opening date                                       | 0.00    | 1.00       | 1.00–1.00|

Significant associations are in bold.

*P < 0.05; **P < 0.001.
an analogue – the Flesch index. Although the Flesch index is an accepted proxy for reading difficulty, actual comprehension in terms of a research study is impossible to capture with a formula. Information sheets are typically augmented by conversations with research or clinical staff at the point of consent. However, the information sheet is what the potential participant can take away from this encounter and if they do not understand it then there are clearly ethical implications. The second limitation concerns our findings relating to recruitment success. We have not conducted a randomised controlled trial, and therefore we cannot rule out confounding. We have tried to capture many different variables that are thought to be important in the largest observational study to date. Further information from qualitative or randomised studies would be helpful.

Policy and ethics implications

Information sheets are approved by ethics committees, and yet they tend to consist of text written at a level far higher than can be assumed of the average reader. Information sheets are also becoming longer, and this may further impede participants’ understanding.

Some have suggested that their length and complexity are increasing because ethics committees and principal investigators now emphasise the legal, rather than communicative, aspects. In our study, the most complex section was often standardised content such as insurance arrangements or confidentiality policies. Other investigators have found that attempts to simplify these template passages are rejected because of concerns of how text alterations might affect the legality of the statements. Clearly this creates an impasse whereby important legal text will never be understood. If this is the case, fear of litigation has distorted the real purpose of study information sheets, leaving review boards protected but participants uninformed.

Principal investigators and ethics committees must balance legal and ethical issues and we consider that this is now out of kilter. We therefore suggest that they critically review their standardised content. It is ethically unsound to allow such text to be included when it is clear that to understand it participants would need their own lawyers to provide an explanation.

Clearly, there is some way to go before information sheets are written to a standard which is likely to be understood by most potential participants. We did find that the information sheets we analysed were considerably easier to read than in some other recent studies, but this provides little solace as they were generally still much more complex than the recommended grade 6 level. Principal investigators and ethics committees must consider the length, complexity, and – most importantly – the purpose of information sheets if this standard is to be achieved.

Acknowledgements

T.W. acknowledges the support the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London, and her NIHR Senior Investigator Award. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Some information in our data-set relates to studies which are still recruiting and so our data-set is not openly available due to confidentiality. However, information regarding the progress of studies supported by the NIHR: CRN Mental Health is available from http://public.ukcrn.org.uk/search/

Appendix

(a) Confidentiality

How will confidentiality and anonymity be ensured?

All your questionnaire responses will be totally confidential: published reports will not allow the responses of any individual worker to be identified in any way, nor will any information about individual responses be fed back to colleagues or managers in your service. The Response Confirmation sheet that you have been given with the questionnaire has a study number on it: only the local research worker knows which number belongs to which staff member, and we have used these numbers only to allow the researcher to check which staff members have returned the questionnaire and which have not.

This section has a Flesch index score of 18.9, which is in the same range as many scientific writings. It is 97 words long.

Will anyone know I’ve taken part?

We won’t tell your manager or anyone else that you have taken part. We also won’t include your name, or anything else which could identify you, in anything we publish.

We will keep your answers separated from your name. We replace names with numbers so that no-one will know what answers you gave.

We have rewritten this section and its Flesch index score is now 73.6 (fairly easy). It is now 53 words long.

(b) Invitation to take part

Invitation to take part in our study

The test we have just carried out has suggested you are suffering from significant health anxiety, excessive concern about your health which is causing you problems. We think we now have ways of helping people with this problem and so you are being invited to take part in what is called a randomised control trial, in which you would be allocated one of two treatments, either a treatment called ‘cognitive behavioural therapy’ which will be given for between 5 and 10 sessions of just under an hour each time, or a simple explanation of what health anxiety is and how it affects people.

This section has a Flesch index score of 28.8, which again falls into the ‘very difficult’ range of scores. It is 103 words long.

Invitation to take part in our study

The test you just did makes us think that you are very worried about your health. We have a new treatment which might help people who worry about their health. The treatment is called cognitive behavioural therapy. We need to know if the treatment works and so we are doing an experiment. In the experiment, we will compare two groups of people. One group will receive cognitive behavioural therapy for 5–10 sessions. Each session will last an hour. The other group will simply receive some information about health worries. The group people are in will be decided at random. Would you like to know more?

Our version has a Flesch index score of 75. It is 107 words long.

This study was funded by the NIHR: CRN Mental Health. The funders had no part in the decision to publish this manuscript.
References

Sense and readability: participant information sheets for research studies
Liam Ennis and Til Wykes

BJP published online September 17, 2015 Access the most recent version at DOI: 10.1192/bjp.bp.114.156687

References
This article cites 0 articles, 0 of which you can access for free at:
http://bjp.rcpsych.org/content/early/2015/09/02/bjp.bp.114.156687#BIBL

Reprints/permissions
To obtain reprints or permission to reproduce material from this paper, please write to permissions@rcpsych.ac.uk

P<P
Published online 2015-09-17T00:05:16-07:00 in advance of the print journal.

You can respond to this article at
/letters/submit/bjprcpsych;bjp.bp.114.156687v1

Downloaded from http://bjp.rcpsych.org/ on January 28, 2018
Published by The Royal College of Psychiatrists

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to The British Journal of Psychiatry go to:
http://bjp.rcpsych.org/site/subscriptions/