

Patient-reported outcome measures in ophthalmology: too difficult to read?

Deanna J Taylor ,¹ Lee Jones ,^{1,2} Laura Edwards ,³ David P Crabb ¹

To cite: Taylor DJ, Jones L, Edwards L, *et al.* Patient-reported outcome measures in ophthalmology: too difficult to read? *BMJ Open Ophthalmology* 2021;**6**:e000693. doi:10.1136/bmjophth-2020-000693

Received 18 December 2020
Accepted 30 April 2021

ABSTRACT

Objective Patient-reported outcome measures (PROMs) are commonly used in clinical trials and research. Yet, in order to be effective, a PROM needs to be understandable to respondents. The aim of this cross-sectional analysis was to assess reading level of PROMs validated for use in common eye conditions.

Methods and analysis Readability measures determine the level of education a person is expected to have attained to be able to read a passage of text; this was calculated using the Flesch-Kincaid Grade Level, FORCAST and Gunning-Fog tests within readability calculations software package Oleander Readability Studio 2012.1. Forty PROMs, previously validated for use in at least one of age-related macular degeneration, glaucoma and/or diabetic retinopathy, were identified for inclusion via a systematic literature search. The American Medical Association (AMA) and National Institutes of Health (NIH) recommend patient materials should not exceed a sixth-grade reading level. Number of PROMs exceeding this level was calculated.

Results Median (IQR) readability scores were 7.9 (5.4–10.5), 9.9 (8.9–10.7) and 8.4 (6.9–11.1) for Flesch-Kincaid Grade Level, FORCAST and Gunning-Fog test, respectively. Depending on metric used, this meant 61% (95% CI 45% to 76%), 100% (95% CI 91% to 100%) and 80% (95% CI 65% to 91%) exceeded the recommended threshold.

Conclusion Most PROMs commonly used in ophthalmology require a higher reading level than that recommended by the AMA and NIH and likely contain questions that are too difficult for many patients to read. Greater care is needed in designing PROMs appropriate for the literacy level of a population.

INTRODUCTION

Health literacy, defined as ‘people’s knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgements and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course’,¹ has become a well-known term in medical and healthcare communications. At its simplest level health literacy is related to assessment of reading and literacy levels required to engage with one’s own health. Studies assessing the reading level of patient educational materials, such as information websites and leaflets, in

Key messages

What is already known about this subject?

▶ Patient-reported outcome measures (PROMs) are being used increasingly to shape policy and practice in healthcare. Previous research has shown that PROMs in other healthcare disciplines have readability levels beyond the average adult. This may limit the population that PROMs represent.

What are the new findings?

▶ The majority of PROMs used in three common eye conditions (age-related macular degeneration, glaucoma and diabetic retinopathy) do not meet public health recommended readability levels.

How might these results change the focus of research or clinical practice?

▶ Greater care is needed in choosing and designing PROMs appropriate for the literacy level of a population.

ophthalmology suggest a majority of patient information is currently not easily readable.^{2,3}

Patient-reported outcome measures (PROMs) are an increasingly used endpoint in clinical trials.^{4,5} A clinical endpoint should be a ‘clinically meaningful measure of how a patient feels, functions or survives’.⁶ In other words, the outcome should be ‘relevant to the patient’.⁷ In ophthalmology, it has been acknowledged that traditional clinical measures, such as visual acuity, do not reflect the patient’s experience or the impact of disease on patients’ lives⁸ and PROMs are often used as outcome measures in ophthalmic clinical trials.^{9–14}

There have been several calls to incorporate PROMs into routine practice.^{5, 15–17} PROMs have been used routinely preoperatively and postoperatively in hernia, hip, knee and varicose vein surgery in the UK.¹⁵ In the field of ophthalmology, they have been piloted before and after cataract surgery in New Zealand, Sweden and the Netherlands.^{15, 18–20} Yet, this incorporation into routine practice is currently the exception rather than the norm.²¹

If PROMs are to be effectively used to shape policy, and manage patients then the



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Optometry and Visual Sciences, City University of London, London, UK

²Institute of Ophthalmology, University College London, London, UK

³Moorfields Eye Hospital NHS Foundation Trust, London, UK

Correspondence to

Dr Deanna J Taylor; deanna.taylor.2@city.ac.uk

instruments designed to elicit them ought to be legible and understandable. Studies analysing PROMs used in orthopaedic surgery,²² oral health²³ and most recently audiology,²⁴ suggest, for example, that the majority of PROMs may have readability levels beyond that of the average adult. Kröll *et al*²⁵ reflect on the parts of society that might be most under-represented as a result of this; indeed, Baker *et al*²⁶ report poorer reading ability to be associated with poorer health. In other words, the people for whom PROMs may be most pertinent, may be the least able to interact with them.

The aim of this study, therefore, was to assess the readability of commonly used PROMs in ophthalmology, and compare these to each other and to population norms for literacy levels.

MATERIALS AND METHODS

Selection of PROMs

A literature search was conducted on 8 August 2018 using search terms related to PROMs and common ophthalmological conditions (“patient reported outcome measure*” or “PROM” or “questionnaire” or “quality of life” or QoL) and [glaucoma or macula or AMD or ARMD or “diabetic retinopathy” or “diabetic macular oedema” or “diabetic macular edema” or “maculopathy”] and [validat* or rasch or develop*]) in order to identify PROMs that had been used in at least one of age-related macular degeneration, glaucoma or diabetic retinopathy to assess quality of life or visual function. These three conditions were chosen because as they are leading causes of blindness both in the UK and worldwide.²⁷ PROMs were screened for eligibility by three independent researchers (DJT, LE, LJ) using Covidence (www.covidence.org), a computerised literature review management software. Literature search results were imported directly into Covidence, where duplicate studies, and studies reporting on duplicate PROMs were removed. The title and abstract for each study were screened by two authors independently, followed by the full-text articles of those deemed to be relevant. Any disagreements at either stage were discussed until consensus was reached. PROMs were excluded if they had not been administered in the English language, or if they were designed for use in children. PROMs assessing knowledge or health beliefs about a condition were also excluded.

Readability measures

Reading comprehension level determines the readability that a text must have so that a reader can understand these written materials; these were calculated using the Flesch-Kincaid Grade Level test, the FORCAST test, and the Gunning-Fog index using readability calculations software package Oleander Readability Studio 2015 (Oleander Software, Vandalia, Ohio, USA).

Flesch-Kincaid Grade Level test

The Flesch-Kincaid Grade Level test²⁸ is one of the most widely used readability measures, with over 3000 citations

in research literature alone. It takes into account both sentence length and syllables per word. Output is in the form of ‘grade level’, the minimum USA grade level which the text is predicted to be suitable for. For example, a Flesch-Kincaid Grade Level score of five would suggest that a text is suitable for those in USA grade 5 (aged 10–11) or higher.

FORCAST test

The FORCAST test,²⁹ initially designed for assessing readability of US military technical reading materials, is considered the most appropriate test for assessing forms, questionnaires and lists because it does not rely on text being in sentence format. Rather, it is calculated taking into account frequency of monosyllabic words in the document. As with Flesch-Kincaid, output is in the form of school grade level.

Gunning-Fog index

The Gunning-Fog index³⁰ is calculated from the average sentence length, and the number of polysyllabic words in a document. Its output is in the form of grade level. This measure was included because of its wide previous use in literature on readability of written healthcare materials.³¹

Data analysis

The American Medical Association (AMA) and the National Institutes of Health (NIH) recommend readability of patient materials should not exceed the reading level equivalent of a child in the sixth-grade (aged 11–12, see table 1).^{32 33} (For context, this is the readability level of J K Rowling’s *Harry Potter and the Philosopher’s Stone*.) Number of PROMs requiring a reading level exceeding this threshold using each readability measure was calculated.

Table 1 USA school grade levels with corresponding typical age group

Grade	Typical age
1st grade	6–7
2nd grade	7–8
3rd grade	8–9
4th grade	9–10
5th grade	10–11
6th grade	11–12
7th grade	12–13
8th grade	13–14
9th grade	14–15
10th grade	15–16
11th grade	16–17
12th grade	17–18

Groups that fall outside 6th grade threshold shaded in grey.

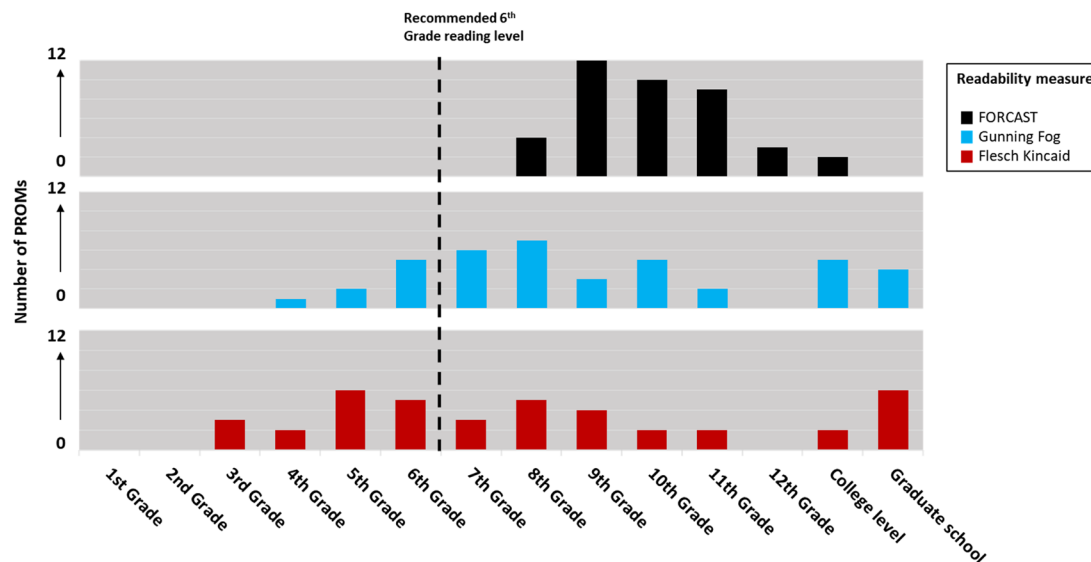


Figure 1 Bar charts showing frequency of patient-reported outcome measures (PROMs) at each grade level, depending on metric used. The black dotted line denotes the 6th grade level; it is recommended that patient materials do not exceed this level.

RESULTS

From a total of 47 PROMs eligible for inclusion based on the literature, 40 were available to access. These were inputted to the Readability Studio in full (including all response options and instructions).

Median (IQR) readability scores were 7.9 (5.4–10.5), 9.9 (8.9–10.7) and 8.4 (6.9–11.1) for the Flesch-Kincaid Grade Level test, the FORCAST test and the Gunning-Fog test, respectively. In other words, the PROMs evaluated would be expected to be understood, on average, by an individual with a reading age of a 13–14, 15–16 or 13–14 years old, respectively. Results remained unchanged when instructions were omitted from all PROMs. Depending on the metric used this meant 61% (95% CI 45% to 76%), 100% (95% CI 91% to 100%) and 80% (95% CI 65% to 91%) fell outside the 6th grade reading level recommended by the AMA and NIH (figure 1).

Characteristics of PROM text

Median word count for PROMs was 416 words (IQR 259–734; minimum words (min) 75, maximum words (max) 4601); 18% of included PROMs had word counts exceeding 1000 words. Documents comprised a median of 10% complex (3+ syllable) words (IQR 8% - 13%; min 4%, max 2%) and a median of 29% long (6+ character) words (IQR 23% - 34%; min 16%, max 50%).

DISCUSSION

An estimated 30 million US adults and 10 million UK adults cannot read beyond a 3rd grade level^{34 35}; almost all PROMs assessed in this study (using any formula) had a readability level beyond this. This is noteworthy. The study's results suggest that most PROM questionnaires commonly used in ophthalmology require a reading comprehension level better than that recommended by

the AMA and NIH for patient material. Moreover, when assessed solely using the FORCAST measure, which is the measure recommended for assessing questionnaire and survey text, all PROMs included in this study fell outside recommended readability levels.

These results support findings from literature in other healthcare disciplines. For example, using the FORCAST measure, PROMs used in oral disease²³ and audiology²⁴ were also consistently found to exceed the recommended 6th grade reading level. Another study (although using a different readability formula) reported the average reading level of 59 PROMs used in orthopaedics to equate to that of 16–18 years old. Beyond the readability scores, several PROMs were also found to be markedly lengthy. Eighteen per cent of PROMs were longer than 1000 words. For context, that is almost half the length of this manuscript. The brevity of PROMs has been highlighted in previous literature as a priority to both ophthalmology patients and clinicians; qualitative findings on important factors in questionnaire design from one study³⁶ includes quotations like 'must be short, practical and useful', 'you need something that's basic and easy to fill in' and 'keeping it obviously as brief as possible'.

As the use of PROMs becomes more widespread, it is crucial that their content is accessible and understandable to the majority of their target population. Missing data in PROMs has been reported as a major problem,³⁷ and it is certainly possible that readability may be a contributing factor to lack of motivation to complete PROMs.³⁸ Prospective confirmation of this with empirical data would be helpful. Indeed, population data has shown that people with lower literacy levels are less likely to participate in volunteer activities and more likely to report poor health than those with high literacy levels.³⁹ The NEI-VFQ 25 is currently the most commonly used

PROM in ophthalmology and had a reading grade level of 10th grade using the FORCAST measure, and 9th grade using both Flesch-Kincaid and Gunning Fog. This is equivalent to the readability of *Moby Dick* by Herman Melville.

The results of this study have several implications for future practice and research. Importantly, PROMs that have already gone through extensive development and validation processes should not simply be discarded. However, when choosing from existing PROMs and when designing new PROMs, substantial attention should be paid to the complexity of language used, particularly with respect to word length. It is worth noting that current recommended practice for PROM design includes input from participants in the form of qualitative investigation, from which items are derived using participants' own language. Therefore, the vocabulary used in PROMs should ideally align with that of a sample of the target population. Yet, the representation of individuals with lower literacy levels in these samples remains an inherent problem. When designing new PROMs, a number of steps can be taken to improve readability. Patients should be involved at each stage of the PROM development process and outreach exercises should be undertaken to ensure that these individuals represent a broad sample of the target population. It is advisable to avoid using technical language, use short sentences, write questions in a conversational style and use words and language consistently.^{24 40 41} Resources such as the Center for Disease Control and Prevention's Plain Language Thesaurus for Health Communications,⁴² and the Living Word Vocabulary⁴³ may be useful for assessing the readability level of certain words, and finding replacements where appropriate. This should be applied to items, response options and instructions. Where use of technical jargon is unavoidable, one may provide a simple glossary of terms used.

The methodology was a key strength of this study. The selection of PROMs was done systematically. A range of validated readability measures relevant to the study's aims were used, each of which having been well-described in the literature, and used in readability studies of PROMs in different disciplines. In addition, this is the first study of its kind in ophthalmology, and highlights an important limitation and factor to consider when using PROMs. Results are limited by the fact that the analysis was restricted to PROMs that had been used in AMD, glaucoma and/or diabetic retinopathy. These conditions were chosen because they are three of the leading causes of blindness both in the UK and worldwide.²⁷ However, future work ought to systematically review readability of PROMs used across ophthalmic conditions and should then concentrate on ensuring PROMs across ophthalmology are at appropriate readability level for their targeted respondents. Furthermore, results are discussed in the context of population norms and general public health guidance but there is no data available on the literacy levels of specific patient populations.

While careful consideration was given to choosing appropriate readability measures to the study's aims, each of these formulae come with their own set of limitations. No readability measure is a perfect measure of comprehension.²³ In addition, other factors, such as formatting, font and font size used, and method of administration, may all impact the final comprehensibility of a PROM. This is a particularly pertinent consideration for PROMs that may be used among a visually impaired population where reading a PROM in its traditional format may not be possible. Readability scores cannot be applied to situations where a PROM is read out loud to the participant; the 'listenability' of a piece of text does not equate to its readability, and listening skills have been recognised as distinct from reading skills.^{44–46} PROMs may be subject to other weaknesses beyond the scope of this study, such as those relating to their psychometric properties, or difficulties establishing unidimensionality. Finally, the AMA and NIH guidelines used as a benchmark reference are based on US literacy levels and may not be appropriate guidelines for literacy levels in other English-speaking countries. While there are no specific standards in the UK for written health materials, the government recommends that public facing written material should not exceed the reading level of a 9 years old (3rd grade).⁴⁷ If this standard were applied to the PROMs identified in this study, 93%–100% of PROMs (depending on readability measure used) would fall outside the threshold.

To summarise, most PROM questionnaires and instruments used in three common eye conditions require a literacy level better than that recommended by the AMA and NIH for patient material. It is likely that a majority of PROMs use language at a level too advanced for most patients to read easily. Greater care is required in choosing and designing PROMs appropriate for the literacy level of a population.

Contributors DJT performed study design, data collection, data analysis and manuscript preparation. LJ and LE performed study design, data collection, data analysis and manuscript critique. DPC performed study design and manuscript critique.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests DPC reports unrestricted grants from Roche, Santen, Allergan; speaker fees from THEA, Bayer, Santen, Allergan; consultancy with Centervue; all outside the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Deanna J Taylor <http://orcid.org/0000-0001-8261-5225>
Lee Jones <http://orcid.org/0000-0002-8030-1211>
Laura Edwards <http://orcid.org/0000-0002-0886-8520>
David P Crabb <http://orcid.org/0000-0001-8754-3902>

REFERENCES

- Sorensen K, Van den Broucke S, Fullam J, *et al.* Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012;12:80.
- Huang G, Fang CH, Agarwal N, *et al.* Assessment of online patient education materials from major ophthalmologic associations. *JAMA Ophthalmol* 2015;133:449–54.
- Edmunds MR, Barry RJ, Denniston AK. Readability assessment of online ophthalmic patient information. *JAMA Ophthalmol* 2013;131:1610–6.
- Vodicka E, Kim K, Devine EB, *et al.* Inclusion of patient-reported outcome measures in registered clinical trials: evidence from ClinicalTrials.gov (2007–2013). *Contemp Clin Trials* 2015;43:1–9.
- Calvert M, Kyte D, Price G, *et al.* Maximising the impact of patient reported outcome assessment for patients and society. *BMJ* 2019;364:k5267.
- Lesko LJ, Atkinson AJ. Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. *Annu Rev Pharmacol Toxicol* 2001;41:347–66.
- Medeiros FA. Biomarkers and surrogate endpoints in glaucoma clinical trials. *Br J Ophthalmol* 2015;99:599–603.
- Denniston AK, Kyte D, Calvert M, *et al.* An introduction to patient-reported outcome measures in ophthalmic research. *Eye* 2014;28:637–45.
- Brown DM, Kaiser PK, Michels M, *et al.* Ranibizumab versus verteporfin for neovascular age-related macular degeneration. *N Engl J Med Overseas Ed* 2006;355:1432–44.
- Rosenfeld PJ, Brown DM, Heier JS, *et al.* Ranibizumab for neovascular age-related macular degeneration. *New England Journal of Medicine*. 2006;355:1419–31. 2006/10/05.
- Krezel AK, Hogg RE, Azuara-Blanco A. Patient-Reported outcomes in randomised controlled trials on age-related macular degeneration. *Br J Ophthalmol* 2015;99:1560–4.
- Chakravarthy U, Harding SP, *et al.* IVAN Study Investigators. Ranibizumab versus bevacizumab to treat neovascular age-related macular degeneration: one-year findings from the IVAN randomized trial. *Ophthalmology* 2012;119:1399–411.
- Gazzard G, Konstantakopoulou E, Garway-Heath D, *et al.* Selective laser trabeculoplasty versus eye drops for first-line treatment of ocular hypertension and glaucoma (LiGHT): a multicentre randomised controlled trial. *Lancet* 2019;393:1505–16.
- Azuara-Blanco A, Burr J, Ramsay C, *et al.* Effectiveness of early lens extraction for the treatment of primary angle-closure glaucoma (EAGLE): a randomised controlled trial. *Lancet* 2016;388:1389–97.
- Devlin NJ, Appleby J. Getting the most out of PROMs 2010.
- Braithwaite T, Calvert M, Gray A, *et al.* The use of patient-reported outcome research in modern ophthalmology: impact on clinical trials and routine clinical practice. *Patient Relat Outcome Meas* 2019;10:9–24.
- Appleby J, Devlin N. Measuring NHS success: can patients' views on health outcomes help to manage performance? *King's Fund* 2005.
- Lundström M, Stenevi U, Thorburn W. Quality of life after first- and second-eye cataract surgery: five-year data collected by the Swedish national cataract register. *J Cataract Refract Surg* 2001;27:1553–9.
- Mollazadegan K, Lundström M. A study of the correlation between patient-reported outcomes and clinical outcomes after cataract surgery in ophthalmic clinics. *Acta Ophthalmol* 2015;93:293–8.
- Stolk-Vos AC, Visser MS, Klijn S, *et al.* Effects of clinical parameters on patient-reported outcome in cataract patients: a multicentre study. *Acta Ophthalmol* 2018;96:586–91.
- Field J, Holmes MM, Newell D. PROMs data: can it be used to make decisions for individual patients? A narrative review. *Patient Relat Outcome Meas* 2019;10:233–41.
- El-Daly I, Ibraheim H, Rajakulendran K, *et al.* Are patient-reported outcome measures in orthopaedics easily read by patients? *Clin Orthop Relat Res* 2016;474:246–55.
- Pace CC, Atcherson SR, Zraick RI. A computer-based readability analysis of patient-reported outcome questionnaires related to oral health quality of life. *Patient Educ Couns* 2012;89:76–81.
- Douglas A, Kelly-Campbell RJ. Readability of patient-reported outcome measures in adult audiology rehabilitation. *Am J Audiol* 2018;27:208–18.
- Kroll T, Wyke S, Jahagirdar D, *et al.* If patient-reported outcome measures are considered key health-care quality indicators, who is excluded from participation? *Health Expect* 2014;17:605–7.
- Baker DW, Parker RM, Williams MV, *et al.* The relationship of patient reading ability to self-reported health and use of health services. *Am J Public Health* 1997;87:1027–30.
- Flaxman SR, Bourne RRA, Resnikoff S, *et al.* Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *Lancet Glob Health* 2017;5:e1221–34.
- Kincaid JP, Fishburne Jr RP, Rogers RL. Derivation of new readability formulas (automated readability index fog count and flesch reading ease formula) for navy enlisted personnel 1975.
- Caylor JS. Methodologies for determining reading requirements of military occupational specialties 1973.
- Gunning R. The technique of clear writing 1952.
- Wang L-W, Miller MJ, Schmitt MR, *et al.* Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Social Adm Pharm* 2013;9:503–16.
- Weiss B. *Health literacy: a manual for clinicians*. Chicago, IL: American Medical Association Foundation, 2003.
- Medicine UNLo. How to write Easy-to-Read health materials: MedlinePlus. Available: <https://medlineplus.gov/etr.html> [Accessed 02 Sep 2019].
- Harding C, Romanou E, Williams J. The 2011 skills for life survey: a survey of literacy numeracy and ICT levels in England 2012.
- Statistics UDoENCFE. Adult literacy in the United States, 2019. Available: <https://nces.ed.gov/datapoints/2019179.asp> [Accessed 05 Jul 2019].
- Somner JEA, Sii F, Bourne RR, *et al.* Moving from PROMs to POEMs for glaucoma care: a qualitative scoping exercise. *Invest Ophthalmol Vis Sci* 2012;53:5940–7.
- Gomes M, Gutacker N, Bojke C, *et al.* Addressing missing data in patient-reported outcome measures (PROMs): implications for the use of PROMs for comparing provider performance 2016;25:515–28.
- Fung CH, Hays RD, RDJQoLR H. Prospects and challenges in using patient-reported outcomes in clinical practice. *Qual Life Res* 2008;17:1297–302.
- OECD. England & Northern Ireland (UK) - Country Note -Survey of Adult Skills first results, 2013. Available: <https://www.oecd.org/skills/piaac/Country%20note%20-%20United%20Kingdom.pdf> [Accessed 05 Jul 2019].
- Badarudeen S, Sabharwal S. Assessing readability of patient education materials: current role in orthopaedics. *Clin Orthop Relat Res* 2010;468:2572–80.
- Adams J, Chapman J, Bradley S, *et al.* Literacy levels required to complete routinely used patient-reported outcome measures in rheumatology. *Rheumatology* 2013;52:460–4.
- Tupper J. *Plain language thesaurus for health communications*. Washington DC: Centers for Disease Control and Prevention, 2009.
- Dale E, O'rourke J. The living word vocabulary the words we know: a national vocabulary inventory 1976.
- Rubin DL. Listenability as a tool for advancing health literacy. *J Health Commun* 2012;17 Suppl 3:176–90.
- Nouri SS, Rudd RE. Health literacy in the "oral exchange": an important element of patient-provider communication. *Patient Educ Couns* 2015;98:565–71.
- Fang IE. The "Easy listening formula". *Journal of Broadcasting* 1966;11:63–8.
- Services GD. Content design: planning, writing and managing content. Available: <https://www.gov.uk/guidance/content-design/writing-for-gov-uk> [Accessed 07 May 2020].