Information Design Journal 16(3), **216–228**

DOI: 10.1075/idj.16.3.05waa

Karel van der Waarde

Measuring the quality of information in medical package leaflets: harmful or helpful?

Situation: Patients, doctors, pharmacists and nurses must receive visual information about medicines. Without instructions, warnings and risk-benefit information, it is not possible to prescribe, dispense, or take medicines appropriately. Research indicates that fifty percent of medicines for chronic illnesses are not taken effectively and that the number of hospital admissions and fatal accidents caused by medicines are significant. The visual design of the information is likely to be strongly related to these statistics. Confusing medical packaging, poor instructions, hard to read package leaflets, and conflicting warnings are commonly found when the visual design of information about medicines is analysed.

Problem: European Union legislation has made user-testing of package leaflets obligatory. These 'readability tests' should guarantee that the text and the design of package leaflets 'enables people to act appropriately'. The EU suggest 'diagnostic testing' as their preferred testing method. This method is well suited to find problems that people have with package leaflets. Unfortunately, the diagnostic test is currently mainly used to provide quantitative data on 'how readable a leaflet is'. Furthermore, it is only obligatory to test package leaflets, and not any of the other information that is necessary to handle medicines appropriately. This kind of testing is therefore unlikely to help alleviate the above-mentioned situation.

Conclusion: Looking at current practice of measuring 'readability' with a diagnostic test, and presenting the results as quantitative data does not do justice to the range of activities that must be supported by well-designed information. There seems to be a need to reconsider the testing process as it is currently used to evaluate the 'readability of package leaflets' in the European Union.

This article is based on a presentation given at the IIID Vision Plus 12 conference 'Achieving measurable results' in Schwarzenberg, Austria in 2006.

The following text provides an outline of a single perspective on the circumstances in 2008. There are many stakeholders involved in the development of information about medicines, and all have different concerns which influence contents and design. A single article cannot do justice to the wide variation of perspectives.

1. Developing information about medicines

In order to explain and clarify some of the issues related to the writing, designing and testing of information about medicines, it is necessary to start with a brief description of the regulatory framework in which information about medicines is developed.

In Europe, medicines are, in most cases, packed in a combination of an inner package and an outer package.

The inner package could, for example, be a blister pack or a bottle and the outer package is usually a cardboard box. Inside this cardboard box, there is a leaflet in the shape of a sheet or a small booklet. All three – inner package, outer package and package leaflet – carry tightly regulated information about a medicine, and all three are written and designed by the pharmaceutical industry. Nink & Schröder (2005) and Raynor (2007) provide more detailed reviews of this process.

The basis for the contents and design of the *package leaflet* is given in a European directive, Directive 2004/27 describes a list of information items that must be mentioned in a leaflet. This directive also states that items must be mentioned in a specific sequence (2004/27, article 59). A leaflet must start with the name of the medicine and end with the date on which the package leaflet was last revised.

The European Medicines Agency (EMEA) – which evaluates and supervises medicines that are sold across Europe – have developed a template as a complement to the directive (QRD, 2008. Human Product Information Templates). This QRD-template adds several statements and is available in 24 languages. The three aims of the template are:

- to make information in package leaflets more consistent across Europe,
- to help the pharmaceutical industry write package leaflets, and
- to make it easier for patients to understand information.

For example, the directive mentions that the package leaflet must include *'the pharmaco-therapeutic group or type of activity in terms easily comprehensible for the patient'*. The QRD-template adds the heading '1. WHAT X IS AND WHAT IT IS USED FOR'. This heading must be placed above the description of the pharmaco-therapeutic group. The 'X' must be replaced by the name of the medicine. This heading is now used across Europe, in all package leaflets, in all languages and for all types of medicines.

In addition to the directive and QRD-template, there are several guidelines related to specific issues. There are, for example, guidelines on the translation of medical terminology into terms that are easier to understand for laymen, and a list of standardized phrases for instructions about storing medicines. There are also guidelines for the visual design, and for the process of assessing the 'readability of the package leaflet' by patients. Some guidelines are provided by the European commission, others by the national regulatory authorities and some by industry associations. The combination of the directive, the template and guidelines form the starting point for the development of the text and visual design of package leaflets.

The information in the package leaflet must be based on the '*Summary of Product Pharacteristics* (SmPC)' in Europe or a similar document applying to another part of the world. These summaries provide a relatively brief overview of information about a medicine – between 10 and 20 pages – and one that is very dense. It is written in a legal version of medical and pharmaceutical jargon. The text is very condensed and certainly not suitable to give to patients. These summaries are the starting point for the text in package leaflets.

There are several general points that must be considered. The directive states that '*the information must be full and comprehensible*'. In other words, it is not allowed to leave information out of the package leaflet: everything that is mentioned in an SmPC must be mentioned in the package leaflet.

Furthermore, the Directive states that *'The package leaflet must be written and designed to be clear and understandable, enabling the users to act appropriately, when necessary with the help of health professionals'* (2004/27, article 63). For information designers, this is an

excellent starting point. It is one of the very few laws that mentions 'design'. The legislation also provides several quality criteria that are related to the readers of the information: 'clear', 'understandable', and 'enabling the users to act appropriately'.

The Directive does not only provide the basis for *writing* and *designing* the package leaflet, but it also stipulates that package leaflets must be *tested*. It states: *'The package leaflet shall reflect the results of consultations with target patient groups to ensure that it is legible, clear and easy to use'* (2004/27, article 59(3)). This shows that the pharmaceutical legislation in Europe is fairly exceptional. Not only does it mention writing and designing, but it also makes some form of usability testing obligatory.

2. Consultations with target patient groups: a readability test

Since a few years, the '*consultations with target patient groups*' are compulsory. Without a 'consultation with patients' or 'assessment by patients' it is not possible to obtain a license to sell medicines in Europe. The law itself does not mention how this assessment needs to be done, and several regulatory authorities have published practical guidelines to indicate which forms of 'consultations' and 'assessments' are acceptable. The British Medicines and Healthcare products Regulatory agency (MHRA) describes the aim of a test as '*The reason for user testing is to help produce a leaflet that most medicine users can use to take safe and accurate decisions about their medicines.*' (MHRA, 2005).

An example of a test method is described in the Readability Guideline (1998), in *Always read the leaflet* (MHRA, 2005) and the Draft Readability Guideline (European Commission, 2006). Alternative test methods are allowed, but these need to prove first that they provide valid results that are comparable to the example method. Very few attempts have been made to achieve this.

For a readability test, it is necessary to have a mock-up of a well-written and well-designed package leaflet. A second essential document is a questionnaire of about 15 questions. These questions should *'Adequately cover any critical safety issues with the medicine'* (European Commission, 2006). Furthermore, test participants need to be invited to a suitable test venue. A quiet, informal space is most convenient.

The test consists of a series of one-to-one interviews. In each interview, a participant reads the test-leaflet and the interviewer asks the questions in the questionnaire. Depending on the requirements of the commissioner, participants can be asked to read and sign a statement relating to confidentiality, or approval of voluntary participation. Interviews can be sound-recorded or video recorded (see Figure 1).

The interviewer takes notes and scores the 'correct answers'. There are two scores per question. The first score indicates if a participant is able to find the right

Figure 1. A participant needs to find information (location) and interpret information (understanding) in a package leaflet during a readability test

location in the leaflet where the answer is mentioned. The second score indicates if the participant can understand the answer. The understanding is checked by asking the participant to rephrase the answer in their own words, or to apply the answer to a specific situation. The interviewer also notes any comments that a participant makes.

A readability test usually consists of three steps: one pilot test of five interviews, a first series of ten interviews and a second series of ten interviews. The aim of the test is to meet the success criteria in a total of 20 participants.

The whole process is described in a 'Readability Report'. This report must include the test details, such as the method used, an explanation on the choice of participants and indicate the language(s) tested. It must also provide the questionnaire, the original and the revised package leaflets. The most important part of the report is the discussion of the answers, of the identified problems and the revisions made to the package leaflet.

3. Testing information: five discussion points

The fundamental assumption of the current regulatory framework seems to be that this form of consultation will help to make package leaflets 'legible, clear and easy to use' and that this in turn this will support patients to take safe and accurate decisions.

The MHRA guidance provides this motivation:

> '[diagnostic testing] is a performance based, flexible development tool which identifies barriers to people's ability to understand and use the information presented, and indicates problem areas which should be rectified. It is particularly useful as part of a leaflet development process.' (MHRA, 2007).

This assumption is rarely questioned or evaluated. Is the diagnostic test used effectively, and do the results really have the required outcomes? In other words, does the selected technique of user testing help *'to write and design package leaflets to make them clear and understandable'*, and does it lead to the ultimate aim of *'enabling users to act appropriately'*?

It must be noted that it is very easy to make critical comments about the suggested processes and the quality of the guidance and templates. Providing relevant information to patients about medicines is a complex challenge which involves a substantial number of stakeholders. A first version of the European legislation was published only sixteen years ago, the European readability guideline is now ten years old, and most of the specific research has been done in the last three years. It is certainly too early to judge particular approaches. However, it seems worthwhile to mention and discuss some of the practical issues.

Below is a list of five considerations about the current practice of readability testing of package leaflets in Europe. These five issues are based on practical experience, interviews, questions by pharmaceutical industries, and a review of the literature.

3.1. Test materials: best practice?

Before a pilot test can commence, it is essential to have a package leaflet that is 'as good as possible'. This writing and design process is supported by several guidelines. Unfortunately, some of the advice is conflicting. Two examples might indicate some of the practical issues.

Example 1:
The QRD-template states that the following sentence must appear at the beginning of a package leaflet:

> *- If any of the side effects gets serious, or if you notice any side effects not listed in this leaflet, please tell your <doctor> <or> <pharmacist>.>*

The following six comments can be made about this sentence.

a. In every readability test interview, every native English speaker stumbles here. The plural of 'side effects' seems to be in conflict with the singular form of 'gets'. After re-reading the sentence, it is either confirmed to be correct English – 'any' is singular – or a remark is made that it should be 'get' and not 'gets'.
b. The sentence appears as the 8th line in a package leaflet. A common reaction is: *"I've only got to here, and they are already talking about side effects. I don't even know what the medicine is for yet."* Line eight is not the correct location to mention side effects because people cannot know what these are at this point in the leaflet.
c. Within the typical patient's vocabulary, any mention of a 'side effect' is serious. Patients should be encouraged to discuss any worry they have with their doctor. Leaving the interpretation of 'serious' to a patient might not be appropriate in all circumstances. Sometimes, 'less serious' effects might be symptoms of very severe side effects.
d. If the package leaflet must 'enable the user to talk about side effects', then this should be investigated. 'How many side effects are mentioned by users, and do we find that appropriate'? Otherwise, there is a direct conflict with article 63(2) of Directive 2004/27 (2004).
e. 'Please' is nearly always redundant, and dilutes the message.
f. The 2006 Draft Readability guideline states: *'Sentences should be no more than about 20 words.'* This 26-word sentence in the QRD-template is in conflict with the Draft Readability guideline and there is no indication which document should prevail.

A writer has the option to modify this particular sentence into a phrase that might be more appropriate. The benefit of modifying a sentence from the template must be balanced against the risk of rejection. The authorities might reject any text that deviates from the QRD-template. Small changes might be acceptable, but this must be discussed on a case-by-case basis.

Example 2:
The MHRA provides the following graphic design advice in a guidance document (MHRA 2008):

– "Keep line spaces clear.
– A column format for the text can help the reader navigate the information.
– Do not use italic fonts and underlining as they make it harder for the reader to recognize the word shape."

This is the official advice of the MHRA to professional designers. (The same advice is also given in the European Draft Readability guideline, 2006). The advice to 'keep line spaces clear' is hard to follow because it is not evident what is meant by 'clear linespace'. The phrase is not mentioned in the typographical literature. The second advice is hard to apply too. There are only very few text formats that do not use columns, such as a ticker tape on television news, but these formats would not be suitable for paper artefacts. Most texts in the latin and greek scripts already use a column format. And where do the authors of this guideline find the evidence that 'italic fonts' make it 'harder for the reader to recognize the word shape'? If that is the case, why have printers and typographers used italics for over five centuries?

Both examples indicate the practical difficulties for writers and designers of package leaflets. The guidelines provide conflicting information and are based on questionable assumptions. The starting point of a readability test for package leaflets is therefore rarely optimal. Furthermore, following the QRD-template and the guidelines leads to package leaflets that are in conflict with the EU-directive which states that: *'The package leaflet must be written and designed to be clear and understandable'*.

3.2. Test results: quantifying data or comments for modifications?

The discussion about the criteria of a readability test touches on very complex issues. Basically, there are two types of results from a readability test. The first type of results are the comments by the test participants that *indicate problem areas or barriers to people's ability to understand and use* the leaflet. These comments can be related to occasions during testing when a location is not found, or information is not understood. They can also be made spontaneously while reading or searching. These responses need to be considered carefully and form the basis for modifications of the text and design.

A second type of result would be a list of correctly found and correctly understood answers. This list does not lead to modifications of the leaflet, but is only used quantitatively to indicate the readability of a package leaflet. There is still a lot of confusion about these quantitative criteria. Beate Beime (2007) points to some of the problems.

The strictest quantitative criteria are currently given by the MHRA. They state: *'Over two rounds of 10 participants on the final proposed leaflet we would expect 16/20 participants to have both found and understood the information'* (MHRA, 2007). A few lines further, it continues: *'Nevertheless, whatever success criteria are proposed, each question must satisfy the criteria individually'* (MHRA, 2007). This is a clear description of a first set of criteria.

In the same document, the MHRA and EU-Draft Readability guideline also describe a second set of criteria. These loosely refer to the criteria used by David Sless and Ruth Shrensky (2006), although this publication is rarely mentioned. The guidelines state that *'90% of literate adults should be able to find the information and of these 90% should be able to understand the information.'*

The publication of both these criteria has lead to several interpretation problems:

- According to the second set of criteria, 90% is an exact measure. In the original text, the words 'at least' are used. The omission of 'at least' makes this criterion hard to apply. It actually states that exactly 10% of literate adults should not be able to find information …
- The guidelines keep mentioning 'two rounds of 10 participants', while the results must be given over 20 interviews. It would be clearer if it stated 'one round of 20 participants'. It is not allowed to make modifications after the first ten interviews, so in practice it is a series of 20 interviews.
- There is a difference between 16/20 and '90% of 90%'. It might only look like 1%, but in practice it means a complete shift in the interpretation of the test results.

According to the 16/20 measure, it is acceptable if 4 people cannot find the answer to a specific question. (Theoretically, 4 people can answer all questions wrongly, as long as the other 16 participants can find all the answers and can indicate that they understood it.) If the 16/20 measure is used, than a test fails if five people make more than one mistake regardless of the question. Furthermore, this criterion does not distinguish any more between 'location' and 'understanding'.

According to the '90% of 90%' measure, all participants could make mistakes. For each question, eighteen people out of twenty must be able to locate the answer. The guidelines do not indicate how the '90% of 18 participants' must be calculated. If the '90% of 90%' measure is strictly used, a test fails if one question is found by 18 participants and correctly answered by 16 participants out of these 18.

The confusion about the different interpretations of the results of the twenty interviews has substantial consequences and is not very practical. A few incorrect answers might make a complete new series of twenty interviews necessary, but shifting from one set of criteria to the other is usually sufficient to exceed the required level.

3.3. Test participants: to modify a leaflet or to reach criteria?

The selection of participants raises some questions too. The MHRA and EU ask to '*Ensure a range of different types of people who are able to imagine needing to use the medicine.*' (2005, p 93; 2006, p 21). In practice, it is difficult to distinguish between 'types of people' if there is no indication how these could be classified.

The MHRA (2005, 2008) suggests to try to include:

> "– people who do not use written documents in their working life
> – people who find written information difficult."

This advice is hard to apply. Especially the second inclusion criterion implies that there are people who do not find 'written information difficult'. Without specifying what kind of information is meant, this guidance is not very helpful.

Schickl (2007) points to another conflict here. She states, about two MHRA publication (2005, 2006):

> The MHRA paper in June 2005 details that there may be special indications, e.g. Alzheimer's disease, where the care-givers may be the appropriate target group for user testing, whereas the additional publication dated June 2006 records that "*…health care professionals and other staff / people who routinely work with medicines information must be excluded to avoid bias…*", which means that care-givers would not represent adequate test persons.

Furthermore, the combination of the criteria and the freedom in selecting test participants lead to a preselection into two groups. The selection depends on the aim of the test. If the aim of the test is to improve the package leaflet, it is useful to involve participants who are more likely to find real problems in a package leaflet. People in this group are – in general – slightly older and slightly less well educated. If the aim of the test is to pass the success scores, it is beneficial to include participants who are more likely to give correct answers. Of course, it is possible to conduct more pilot tests to keep improving the leaflet, but if the focus is mainly on 'the scores', there is little incentive to continuously improve package leaflets.

3.4. Guidelines: prescriptive or performance?

There is a fundamental conflict in the current framework between regulations that are prescriptive and regulations that are performance based. These two approaches are very hard to reconcile. One the one hand, it is essential to take the comments of test-participants into account and base modifications on these comments. On the other hand, it is necessary to follow guidelines telling exactly what kind of information needs to be mentioned and where and how this should be presented. The 1998 Readability guide realized this inconsistency and stated:

> 'It may therefore be acceptable for a package leaflet, which achieves an acceptable level of performance in a readability test, to deviate from the guidance in this section.'

This phrase is not included in the 2006 Draft Readability Guideline (2006). The MHRA refers to a conflict between the results of readability tests and the use of the QRD-template (2008):

> 'Findings from the survey indicate that the wordings of many of the headings and subheadings in the QRD template are not well understood by patients. You should make sure that when preparing your mock-up leaflet for testing you reword your leaflet to ensure that all the information is translated into terms which the patient can understand.'

Although the above phrases indicate that the problem is recognized, it is far from being resolved. Every readability test will highlight the conflicts between regulations, guidelines and QRD-templates.

The two examples provided in Section 3.1 display this conflict, too. Test participants are likely to stumble over the obligatory phrase *'If any of the side effects gets serious'*. Despite of the MHRA advice, many pharmaceutical industries are very hesitant to change the wording. The second example illustrates the conflict too. If none of the test participants has had any problems with the reading or interpretation of *italic type*, but the guideline suggest that this '*makes it harder for the reader to recognize the word shape*', then which information should be followed?

A second fundamental issue in relation to the guidelines is that none of the guidelines mentions how the answers of participants should be related to the modifications. The MHRA stated: *'If testing reveals barriers to understanding, carefully considered changes to the leaflet will be needed to improve it.'* (MHRA, 2005). Unfortunately, guidelines do not indicate **how** to respond to a problem. What needs to be changed? The wording? The layout? The text size? Would it be useful to add an illustration or a pictogram? None of the guidelines addresses how the comments of test participants might need to be related to changes in the text or design of a package leaflet. Furthermore, it is not clear **when** it is necessary to modify the package leaflet. Should a leaflet be changed if the interviewer considers it a 'sensible remark', even if it is only given by one participant? Or do at least two participants have to mention the same thing?

A further practical complicating factor related to the guidelines about testing is that the European guideline (Draft Guideline, 2006) is still in a 'draft phase'. This draft is strongly based on the British Committee on Safety of Medicines Working Group on Patient Information. This group published "Always read the leaflet" in 2005 and many of the recommendations have become full guidelines within the UK. The guideline is already in force in the United Kingdom, but very similar guidelines are still in a 'draft-phase' in Europe. In practice, this means that the British guidelines are followed as closely as possible, but for applications in other EU countries, of for European applications, this is not based on enforced EU-guidelines.

3.5. The results: are the results optimally used?

The difference between 'comments for improvements' and 'scoring to reach the threshold' (see section 3.2) has an undesirable consequence. In practice, it means that the pilot test is mainly undertaken to retrieve the comments for improvements and the leaflet is modified on the basis of these comments. The '20-people test' is conducted to show that a package leaflet reaches the required level. The guidelines imply that it is not allowed to make any changes to the leaflet or questionnaire during this test.

However, it is very likely that this test will also reveal comments that – if integrated – would improve the leaflet. Unfortunately, the comments in the '20-people test' that could lead to improvements need to be ignored. The integration of these comments into the leaflet would require another twenty interviews. It is even likely that these comments are not even mentioned in the Readability reports. Although they would improve the leaflet, it might provide an argument for the authorities to demand another round of testing. This risk is simply too high in practice. The focus on the score means that the suggestions for improvements offered by twenty test participants must be ignored.

Although the increasing experience with readability testing leads to regular modifications of the guidelines, it does not lead to a modification of the QRD-template. In practice this means that every test will reveal the same issues over and over again.

A third indication that the results are not being optimally used is that the readability reports are not publicly available. In contrast with the scientific data of the clinical trials, the reports of the readability test are not disclosed. This prevents an improvement cycle based on learning from each other.

3.6. More issues?

The abovementioned five issues are probably the most prominent ones. There are a range of others that need attention too. Five examples are:

- Readability testing reports are required in English only. It is therefore common practice to test package leaflets in English only. This avoids the costs of translating the leaflet and translation of the verbatim comments of test participants. The practical consequence is that patients in all other European languages will receive a translated text that was tested in an English context. This assumes that 'all European patients react similarly to British patients'. It is not clear if this assumption is correct.
- The test method is based on work related to Australian Consumer Medicines information (CMI) (Sless, 2006). These CMIs are printed and provided by pharmacists. Additional oral information is therefore likely to be available. This is fundamentally different from the European situation where a patient can only see the leaflet after the box is opened and oral advice is not available. Ignoring the availability of advice from pharmacists might need to be reconsidered.
- The selection of the most relevant questions for the questionnaire is now based on a perception of the 'key messages for safe use'. This selection is not based on actual risks or on actual problems that people have when they handle medicines. It is not necessary to provide evidence that the question really are the 'key messages'.
- Medicines that are used only by health professionals – such as in operating rooms or hospital wards – must have a package leaflet too. It is highly unlikely that a patient ever receives a leaflet for these medicines and the investment in making these leaflets understandable for patients might need to be diverted to improving the instructions for health professionals.
- It is likely that the text and design of a package leaflet will be changed after a readability test was conducted. It is common that modifications are made during the registration process and that new scientific information is added to the Summary of Product Characteristics. These modifications usually do not lead to a new test, but it has the effect that the 'readability scores' for the approved version of the leaflet are not known.

4. Discussion: What can we learn?

The previous section presents several discussion points related to current readability testing practice. It is clear that testing is a valuable step in the process of developing package leaflets that are suitable for patients. The 'good examples' on the MHRA website indicate that substantial progress has been made. Furthermore, the publication of regulations and guidelines has positively fuelled the discussions about the value of printed information for patients. Within a few decades, the focus has shifted from 'is it really necessary?' to 'what is the most appropriate way of providing information about medicines?'.

However, it is also clear that testing in practice as it is undertaken at the moment does not optimally improve package leaflets. Although the results of pilot tests clearly improve the writing and design of package leaflets, the series of twenty interviews does not add much to the quality of the leaflets.

As a first step, it seems worthwhile to reconsider the discussion points from Section 3 and transform these into an action list:

a. Package leaflets must be well designed and well written. This can be 'evidence based' or based on 'best practice'. Guidelines and templates that impede the development of appropriate leaflets need to be modified into 'performance based guidelines'.
b. The test criteria need to be reconsidered and clarified.
c. The selection of test participants needs to be clarified.
d. The guidelines need to be reconsidered and need to become 'evidence based'.
e. The results of readability tests need to be optimally used, not only to improve package leaflets, but also to improve guidelines, templates and regulations.

Despite all the efforts related to the improvement of package leaflets, it is still not known if these efforts bring us any further in achieving the aims as they are described in the legislation: to enable users to act appropriately (EU directive, 2004), and to take safe and accurate decisions about their medicines (MHRA, 2005).

In order to find out if package leaflets really contribute to these aims, it seems essential to broaden the scope of the development process. It is unlikely that package leaflets on their own could be able to achieve the general aims. It is necessary to look at both the wider process as well as at more artefacts. The additional discussion points mentioned in Section 3.6 seem to point to this wider scope, too. The following three points might need to be considered.

4.1 Look at all visual artefacts in a specific context

One of the fundamental practical issues is that patients do not look at a package leaflet only. Package leaflets are always only one artefact and are seen together with the inner and outer packaging. And even this tightly regulated combination of outer packaging, inner packaging and package leaflet is modified during the dispensing and administration of medicines.

For example, additional labels are added to the outer packaging of medicines that are acquired in a local pharmacy. These labels are far from standardised across Europe. The pharmacist labels might for example contain warnings, instructions, details of the patient, names of the prescribing doctor and pharmacist, insurance details, and dates. These labels are not included in the test, and it is unlikely that the information is directly related to the information in the package leaflet. The example in Figure 2 shows the handwritten instructions of a pharmacist. It is unlikely that a package leaflet would help 'to make safe and accurate decisions' about the correct dose in this particular situation.

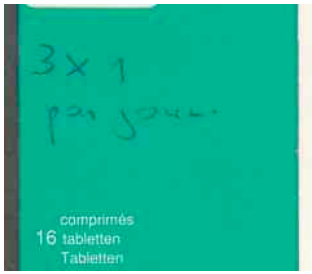

Figure 2. Handwritten instructions on outer packaging is still very common. This dosage instruction – in French – states: '3 x 1 per day'. Does this mean '3 tablets per day: one in the morning, one in the afternoon, and 1 in the evening? or does it mean '1 tablet per day for three consecutive days: one today, one tomorrow and one the day after'? It is the difference between: '(3x1) per day' or '3 x (1 per day)'.

A second situation is the use of medicines in hospitals. In most situations, the hospital pharmacy will have removed the outer packaging and package leaflet before sending the medicine to the appropriate ward. Even the inner packaging might have changed, for example when a blisterpack has been cut to provide single doses or when a syringe needed to be prepared.

In order to provide suitable information, it is necessary to consider what people really get to see when they

handle medicines. This is not only the package leaflet, but usually a combination of several leaflets, boxes, blister-packs, bottles and labels. Medicines are used in different contexts, different ways of administration and by different people – not only by patients, but also pharmacists, nurses and doctors. Ignoring these differences makes it unlikely that appropriate information can be developed.

Furthermore, it is essential to make use of digital technology to provide personalized information. Taylor-ing information is not only possible, but patients simply expect it. At the moment, there is very little progress on the relation between information on paper and the complementing digital resources. It is clear that a combination of information on paper and in digital format might benefit patients.

4.2 Integrate in process: writing, designing and writing

In practical information design work, there is a real imperative to produce measurable performance levels of usability. The current development process for information about medicines does not seem to support this. Readability testing is undertaken at the end of the writing and designing process and every test starts from the assumption that a patient does not know anything about a medicine. Any previous experience that participants might have is not taken into account and is actively ignored. It is common in interviews to try to focus the attention of participants back to the package leaflet by asking 'What does this particular leaflet tell you to do?' This implies that the answer might be different from existing opinions. This results in answers that are correct according to the questionnaire, but it is unlikely that this is a 'externally valid' approach. The results would be far more reliable if they are compared to the results of a 'benchmark test'. A benchmark test needs to be done before any development starts. Such a test would indicate the current knowledge and likely behaviour of people in particular situations. The information could build on this. A comparison of the answer during a readability test with the results of a benchmark test would reveal if a leaflet is really succesful in changing opinions.

4.3 Start from the highest risks

The points mentioned in sections 4.2 and 4.3 seem to suggest that everything needs to be tested, but this is not the case, nor would it be very practical. In order to achieve the aims of the EU and MHRA, it might be beneficial to consider an alternative approach in which the 'actions with the highest risks' are investigated first. For example, recent research for Lloydspharmacy (2008) indicates that about one in five patients admit that they took prescription medicines incorrectly. Further investigations are necessary to find out what the highest risks in these cases are. It is likely that this will reveal that there are very fundamental differences between medicines. These differences are currently ignored in the information that is given to patients. It might be worthwhile to differentiate the visual information according to the potential risks involved.

Based on this discussion, it seems that three steps need to be taken. The first step should be to make package leaflets as good as possible within the current regulatory framework. At the same time, it is possible to identify barriers and indicate problem areas which should be rectified. The second step should be to rectify the problems and to consider the context in which medicines are used. Accurate data about the use of different types of medicines, together with their actual risks, costs and error rates, need to be collected. These data should form the basis for 'evidence based guidelines for the development of performance based information about medicines'. The third step is to develop information that responds to actual differences.

It is unlikely that a package leaflet will be 'the most optimal format in all circumstances'. A variety of alternative formats that are sensitive to the differences between medicines, people, languages, and contexts of use need to be developed. The results of the current readability tests already show the need for this variety, but the regulatory framework prevents enthousiastic developments of alternatives.

5. Conclusion: helpful or harmful?

It is clear that user testing has increased the awareness of the role of patients in the information development process. This has resulted in many package leaflets that are evidently better than their predecessors. It is easier for patients to find information and it is easier to understand this information. Furthermore, the translation of medical jargon into understandable terminology is a substantial improvement. This can all be seen as 'helpful'.

However, these improvements are all on a very modest scale. It does not tell us if 'users are enabled to act appropriately'. A readability test only looks at one specific artefact. It does not look at the process of medicine taking, nor at the process of 'learning about medicines'. A readability test will never reveal if 'a package leaflet serves to ensure that medicines are used both safely and appropriately.' This can be seen as 'harmful'. It's harmful because it gives the false impression that 'we've done what we could'.

If the real aim is to 'enable users to act appropriately', another approach is required. This approach would need to start from the actions that need to be performed safely and effectively.

Unfortunately, it is not possible to develop practical alternatives if the current directives, guidelines and templates are used. Any alternative approach will be 'in addition to' a package leaflet and is likely to be 'illegal'. There is a fundamental choice between "is it necessary to improve the information about medicines for patients?" or "is it enough to provide a 'one size fits all' standard?".

References

Beime, B., Lawrence, M., Rieke, C. (2007). User testing. The practical application of readability user tests in national and international marketing authorisation procedures. *Regulatory Rapporteur (TOPRA)*, May 2007, pp. 8–13.

Directive 2004/27/EC of the European Parliament and of the Council of 31 March 2004, amending Directive 2001/83/EC on the Community code relating to medicinal products for human use, 31 March 2004

European Commission. (1998). *Guideline on the readability of the label and package leaflet of medicinal product for human use*, available from: http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/vol-2/c/gl981002.pdf

European Commission. (2006). *Draft Guideline on the Readability of the Label and Package Leaflet of Medicinal Products for Human Use, Revision September 2006*, available from: http://ec.europa.eu/enterprise/pharmaceuticals/pharmacos/docs/doc2006/09_2006/readability_consultation_2006_09_25.pdf

Lloydspharmacy. (2008). *More than eight million patients admit medicine mistakes*, Pressrelease August 10, 2008. http://www.lloydspharmacy.com/wps/wcm/connect/a410b8004ad5e-848b3e3f7ee9b77d65b/AUG08_MUR.pdf?MOD=AJPERES

MHRA. (2005). *Guidance on the user testing of patient Information leaflets*. MHRA, June 2005.

MHRA (Committee on Safety of Medicines. Working Group on Patient Information) 2005. *Always read the leaflet. Getting the best information with every medicine*. London: The Stationary Office.

MHRA. (June 2006). *Additional Qs and As in relation to the new legal requirement to undertake consultation with target patient groups (Compliance with Article 59(3) of Council Directive 2001/83/EC)*.

MHRA. (March 2007). *Further guidance on designing patient information Leaflets and how to achieve success in user testing. MHRA*, Patient Information Quality Unit.

MHRA. (April 2008). *Questions and Answers on PLPI PIL user testing*. PLPI Unit.

Nink, Katrin & Schröder, Helmut. (2005). *Zu Risiken und Nebenwirkungen: Lesen Sie die Packungbeilage?* Bonn: Wisenschaftliches Institut der AOK. WidO 53.

Raynor, D., Blenkinsopp, A., Knapp, P., Grime, J,, Nicolson, DJ., Pollock, K., Dorer,G., Gilbody., S, Dickinson, D., Maule, AJ. & Spoor, P. (2007). *A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines (HTA Programme Report).* Southampton: National Institute for Health Research.

QRD. (2008). Human Product Information Templates, adopted by the QRD Working Group, http://www.emea.europa.eu/htms/human/qrd/qrdtemplate.htm

Schickel, Ursula. (2007). *The impact of the revised EU-legislation concerning particular needs of blind and partially sighted patients and user testing – a challenge for pharmaceutical companies with focus on Germany*. MA-thesis. Bonn: der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Sless, D., Shrensky, R. (2006). *Writing about medicines for people.* Communication Research Institute. http://www.communication.org.au/

About the author

Karel van der Waarde studied graphic design in the Netherlands (Eindhoven) and in the UK (Leicester, Reading). He received his doctorate in 1994 for a dissertation entitled: 'An investigation into the suitability of the graphic presentation of patient package inserts'. In 1995, he started a design research consultancy in Belgium specializing in the testing of information design. Most of the projects are related to pharmaceutical information. His company develops patient information leaflets, instructions, forms, protocols, and the information architecture for websites. Karel van der Waarde is moderator of the InfoDesign and InfoDesign-cafe discussion lists and professor in Visual Rhetoric at AKV|St Joost, Avans University, Breda (The Netherlands). e: waarde@glo.be