

University of Groningen

## Patient perspectives in the benefit-risk evaluation of drugs

de Vries, Sietske

DOI:

[10.1016/j.jpsychores.2013.11.003](https://doi.org/10.1016/j.jpsychores.2013.11.003)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2015

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

de Vries, S. (2015). *Patient perspectives in the benefit-risk evaluation of drugs*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.1016/j.jpsychores.2013.11.003>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Chapter 2

## Construct and concurrent validity of a patient-reported adverse drug event questionnaire: a cross-sectional study

Sieta T. de Vries<sup>1</sup>

Flora M. Haaijer-Ruskamp<sup>1</sup>

Dick de Zeeuw<sup>1</sup>

Petra Denig<sup>1</sup>

Health and Quality of Life Outcomes 2014;12:103.

<sup>1</sup> Department of Clinical Pharmacy and Pharmacology, University of Groningen,  
University Medical Center Groningen, Groningen, The Netherlands

## **Abstract**

### *Background*

Direct patient-reported information about adverse drug events (ADEs) is important since it adds to healthcare professional-reported information about the safety of drugs. Previously, we developed an instrument to assess patient-reported ADEs in research settings. The aim of this study is to assess the construct and concurrent validity of the questionnaire.

### *Methods*

Patients on at least an oral glucose-lowering drug completed the ADE questionnaire, the World Health Organization Quality of Life-BREF, and the Treatment Satisfaction Questionnaire for Medication (TSQM). The ADE questionnaire assesses ADEs for any drug that the patient uses. Construct validity was assessed by testing whether patients reporting an ADE had a lower general quality of life and physical health than those not reporting an ADE, using Mann-Whitney U tests and t-tests (significance level  $<0.05$ ). For concurrent validity, we tested whether ADEs that patients associate with particular drugs in the ADE questionnaire are documented in the Summary of Product Characteristics (SPC) of those drugs, and whether patients who report an ADE with the use of metformin on the TSQM, mention metformin as a drug associated with an ADE on the ADE questionnaire. Agreement of 70% with the SPC was considered satisfactory. Sensitivity and positive predictive value (PPV) were calculated for the comparison with the TSQM, where 70% was used as the cut-off level for sufficient concurrent validity.

### *Results*

We included 135 patients (mean age 64 years, 35% women). Patients who reported an ADE ( $N = 37$ ) had a lower general quality of life and physical health than those not reporting an ADE ( $P < 0.05$ ). For 78 of the 146 reported ADEs (53%), patients mentioned at least 1 particular drug associated with the ADE. After clustering related ADEs, this resulted in 56 patient-reported ADE-drug associations. Of these, 41 (73%) were in agreement with information in the SPC. Finally, the questionnaire had a sensitivity of 38% and PPV of 79% for assessing ADEs associated with metformin.

### *Conclusions*

The construct validity of the patient-reported ADE questionnaire was sufficient for reporting any versus no ADE, but the concurrent validity was only partly demonstrated. Therefore, the questionnaire needs to be adapted before it can be used.

## Introduction

The safety of a drug is monitored and assessed in clinical trials and observational studies [1,2]. Currently, the attribution of adverse events to a drug and the assessment of the severity of adverse drug events (ADEs) in research settings is primarily conducted by healthcare professionals [3]. It has, however, been shown that healthcare professionals downgrade the severity of ADEs experienced by patients [4]. Additionally, it has been shown that healthcare professionals underestimate symptomatic, subjective ADEs [4-7]. In a literature review, it was for instance shown that ADE rates of constipation with the use of the glucose-lowering drug metformin ranged from 0.6-1.0% when reported by healthcare professionals, and was 21% when reported by patients [6]. Therefore, regulatory authorities acknowledge the added value of patient-reported outcome instruments [8,9] in which the patient is the direct source of information [8,10]. This acknowledgement is especially the case for many symptomatic ADEs for which there is no objective test. Assessment of such ADEs is important, since they influence a patient's quality of life (QOL) [7]. Previous studies showed that an increase in total scores of the number, frequency, and severity of experienced ADEs is associated with a decrease in QOL [11,12]. In addition, patients who report an ADE have a lower general health perception than patients who do not report an ADE [13].

Although some patient-reported instruments to assess ADEs exist (e.g. [14-16]), a generic instrument not limited to a specific ADE, or drug, and including questions about the nature (e.g. frequency, severity) and causality is not available. Therefore, we previously developed such an instrument (**Appendix 1**) [17]. This patient-reported ADE questionnaire is generic, checklist-based, includes questions about the nature and causality of the ADE, and is intended for research purposes in clinical trials and observational studies. The content validity of the instrument has been established and was adequate (**Chapter 1**) [17]. Further validation is needed, in particular given reported concerns about the validity of patient-reported ADEs (e.g. incorrect attributions of symptoms to drugs) [5].

The aim of the current study is to assess the construct and concurrent validity of the patient-reported ADE questionnaire. For the construct validity, the association between patient-reported ADEs and QOL is tested. With respect to the concurrent validity, the focus is on 1) concurrence between reported ADE-drug associations and known ADEs of those drugs, and on 2) agreement between ADE-drug reporting in the generic ADE questionnaire and a treatment/drug-specific questionnaire with a differently phrased question. The results of this study will help to establish whether or not the patient-reported ADE questionnaire is sufficiently valid to investigate experienced ADEs in clinical trials and observational studies, and how to further improve the questionnaire. More information about patient-reported ADEs will increase the knowledge about the

safety of drugs, which in the end can be used in the decision to approve or disapprove the drug to the market and in clinical practice.

## **Methods**

This study had a cross-sectional design and is part of a larger study about the development and validation of the patient-reported ADE questionnaire. For pragmatic reasons, we focused on patients with type 2 diabetes, who may be expected to use a variety of drugs with associated ADEs. Patients prescribed oral glucose-lowering drugs were included in this study since that is a valid proxy for type 2 diabetes. Previously, we have reported on the development process of the ADE questionnaire and assessment of its content validity and reliability (**Chapter 1**)<sup>[17]</sup>. We now present the follow-up study, assessing the construct and concurrent validity of this questionnaire. For this part, additional data were collected during the second measurement of the previous study, assessing the patients' QOL and experiences with metformin, a drug which was expected to be used by most of the included patients (see below). The study was carried out in accordance with the Code of Ethics of the World Medication Association (Declaration of Helsinki) for experiments involving humans. The Medical Ethics Committee of the University Medical Center Groningen (METc UMCG) in the Netherlands determined that ethical approval was not needed for this study (reference number M12.112446).

### **Participants, procedure and data collection**

Patients who were aged 18 years or older and had been dispensed at least an oral glucose-lowering drug were recruited in 2012 via pharmacists in the northern part of the Netherlands, including two pharmacies in a village and two in a town. The pharmacists selected patients who fulfilled the inclusion criteria from their database and sent an information letter with consent form to these patients. One of the researchers (STdV) contacted pharmacists until the number of consenting patients was around 150. Those patients who had an e-mail address, were able to access the internet, and gave informed consent completed the ADE questionnaire twice with a one week period in between. Patient characteristics reported in the current study were based on patient-reported information collected at the second measurement. The pharmacists provided prescription data covering a 1 year period for all patients. They did this at one point in time, after most patients had completed the second measurement. Since chronic medication is prescribed for 3 months in the Netherlands, the most recent prescriptions between 3 months before and up to the date of questionnaire completion were used to describe the medication prescribed to the patients.

## Measures

### *Patient-reported ADEs, ADE-drug associations, and causality score*

The patient-reported ADE questionnaire is a generic questionnaire and contains a checklist with 252 symptoms, categorized in 16 body categories, which patients can check as being a potential ADE for any of the drugs they use (**Appendix 1**)<sup>[17]</sup>. Fourteen additional questions about its nature are asked for each potential ADE, including questions about causality and causal drugs (e.g. “Of which drug or drugs do you think this side effect is the result?”). Branching was used for these additional questions, which means that they were only presented if a symptom was checked as being a potential ADE. Each included symptom in the checklist is linked to a Lowest Level Term of the Medical Dictionary for Regulatory Activities (MedDRA®) terminology version 13.0. For the current study, a web-based version of the questionnaire was used which was constructed using the Unipark Enterprise Feedback Suite 8.0 version 1.1<sup>[18]</sup>. Patients could mention one or more than one drug to be associated with one ADE in the questionnaire, for example, increased stool frequency caused by metformin and tolbutamide. Also, patients could mention the same drug or drugs for multiple ADEs, for example decreased weight and abdominal discomfort caused by liraglutide. Furthermore, patients could check multiple, related ADEs describing one overall ADE, for example, diarrhea and fecal incontinence, which they associated with the same drug or drugs. These related ADEs were clustered by two researchers (STdV and PD) independently<sup>[17]</sup>. This resulted in four possible ADE-drug associations: single ADE-single drug, single ADE-multiple drugs, clustered ADE-single drug, clustered ADE-multiple drugs.

### *Quality of life*

QOL was assessed using the Dutch version of the World Health Organization Quality of Life-BREF (WHOQOL-BREF) questionnaire<sup>[19]</sup>. The WHOQOL-BREF measures a patient’s QOL and contains 26 items, including one item assessing general health, and one item assessing general well-being. These two general items were summed resulting in a general QOL score. The other items comprise four domains of QOL, namely physical health, psychological health, social relationships, and environment. An association between the ADE questionnaire and the general QOL score and physical health domain were used as indicating construct validity of the ADE questionnaire. The WHOQOL-BREF is a reliable and valid questionnaire to measure QOL<sup>[20]</sup>. In our study, the internal consistency ranged from 0.651 (domain social relationships) to 0.836 (domain psychological health) and was 0.806 for the physical health domain.

### *Known ADEs*

Although there is no gold standard for which ADEs are ‘true’ ADEs, the ADEs that are documented in the Summary of Product Characteristics (SPC) can be considered as known ADEs of a drug. Two researchers (STdV and PD) independently determined whether all the patient-reported ADE-drug associations (of both the single and clustered ADEs) could be confirmed with the information in the SPC as indicating concurrent validity. Dissimilarities between the two researchers were resolved by discussion. The SPCs of the reported drugs were retrieved from the website of the Dutch Medicines Evaluation Board <sup>[21]</sup>.

The comparison of ADE-drug associations with the information in the SPC was additionally assessed by taking into account a patient-reported causality assessment. We developed a patient-reported causality scoring system using the additional questions per potential ADE in the questionnaire. This scoring system was based on the items of the Naranjo causality classification <sup>[22]</sup>, and patient-reported aspects of causality assessment <sup>[23]</sup>, and includes the following items of the patient-reported ADE questionnaire: the question why the patient links the symptom to drug use with the option to check more than one of the answer categories assessing causality, the question which drug the patient associates with the ADE, how certain the patient is about this ADE-drug association, and the yes-no question assessing whether or not the patient can think of other reasons for experiencing the ADE. Scores could range from -2 to +10 (**Appendix 3**; supplemental table 1).

### *Treatment/drug-specific ADEs*

The Treatment Satisfaction Questionnaire for Medication (TSQM) <sup>[24]</sup> was applied for the use of metformin. Metformin is the first-choice initial glucose-lowering drug in the treatment of patients with diabetes <sup>[25,26]</sup>, and likely to be used by most of the patients included in our study. Although the TSQM aims to assess treatment satisfaction, the questionnaire contains four items about side effects. One of these items asks for whether or not any ADE was experienced as a result of taking the drug in question (“Do you experience any side effects of the use of metformin?”). The concurrent validity of the ADE questionnaire was examined by using the answers given to this question as the gold standard and comparing them with the reporting of metformin as a particular drug associated with an ADE in the ADE questionnaire. Of note, no recall period is specified for this question in the TSQM. Patients who reported that they had used metformin in the previous 4 weeks had to complete the TSQM.

## Analyses

The QOL of patients who reported one or more ADEs and those who did not report an ADE in the patient-reported ADE questionnaire was expressed with mean and median scores and differences were tested using Mann-Whitney U tests (for non-normally distributed variables, observed for general QOL in the study population, where kurtosis was  $>1$ ) and t-tests (physical health, psychological health, social relationships, and environment domains). Significance levels were used to indicate sufficient validity, and effect-size of the difference in QOL was calculated to indicate clinical relevance. We calculated effect-size as the (mean QOL of those reporting  $\geq 1$  ADEs - mean QOL of those not reporting an ADE) / standard deviation of the QOL scores of those not reporting an ADE. Moderate (0.5 - 0.79) or large ( $>0.8$ ) effect-sizes were interpreted as clinically relevant [27].

Differences in patient characteristics between those who mentioned at least one particular drug for any of their ADEs and those who did not mention a particular drug for any of their ADEs were tested. In addition, differences in ADE characteristics between the ADEs for which patients mentioned at least one particular drug and those ADEs for which patients did not mention a particular drug were tested. T-tests, Fisher-Freeman-Halton tests, and Pearson  $\chi^2$ -tests were used for these analyses.

The percentage of agreement was used in the comparison between the patient-reported ADE-drug association and the information in the SPC of the drug. The percentage of agreement was calculated at the ADE level. This calculation means that in case multiple drugs were reported for one ADE, the confirmation that the ADE was acknowledged in the SPC of at least one of the drugs was considered as agreement in the analyses. We did not expect full agreement between the patient-reported ADE-drug associations and the information in the SPC, since it has been demonstrated that some ADEs are lacking in the SPC [28,29] and that patient reports may provide additional ADEs to those noted in the SPC [5,30]. Therefore, we used a cut-off level of at least 70% agreement between the patient-reported ADE-drug associations and known ADEs as indicating sufficient validity. A sensitivity analysis was performed in which the patient-reported causality score was taken into account. In this analysis, only the cases indicating a single ADE-single drug association could be included, since the causality score was measured at this level. We assessed the agreement for all such cases, and for the cases with a causality score higher than or equal to 1) the median of the causality scores, and 2) the third quartile.

For the comparison of the patient-reported ADE-metformin association in the ADE questionnaire and the TSQM, a positive outcome was defined as the detection of an ADE with the use of metformin. The number of true positives, true negatives, false positives, and false negatives are presented. These numbers were used to calculate the sensitivity



<sup>[31]</sup> and positive predictive value <sup>[32]</sup>. A cut-off level of 70% was used as indicating sufficient sensitivity and/or positive predictive value. Most of the patients on a drug are expected not to experience an ADE which will result in a high number of true negatives. Therefore, specificity and negative predictive value were not calculated.

The analyses were conducted using IBM SPSS Statistics version 20 (Armonk, New York, USA) and P-values of <0.05 were considered statistically significant. Exact confidence intervals based on binomial probabilities were calculated for the validity measures <sup>[33]</sup>, using Stata version 12 (Stata Corp., College Station, TX).

## Results

The number of patients included in this study was 135. The mean age of these patients was 64 years (SD: 9), and most of them were male (65%) (**Table 2.1**). Metformin was the most commonly prescribed oral glucose-lowering drug (94%) and the median number of total systemic drugs prescribed was 7 (interquartile range: 5–9). Thirty-seven patients (27%) reported at least one ADE, with a total of 146 ADEs (median: 3, range: 1–11). In 58% of the cases, the ADE had started more than 12 months ago. Most of the reported ADEs had not yet improved or disappeared (82%).

### Construct validity

Patients who reported one or more ADEs on the questionnaire had a lower self-reported QOL than patients who did not report an ADE (**Table 2.2**). This difference was statistically significant for physical health ( $P < 0.01$ ), and the general QOL score ( $P < 0.05$ ), and turned out to be clinically relevant ( $d = 0.53$  for the general QOL score and  $d = 0.50$  for physical health). Although not statistically significant, patients who reported more than one ADE ( $N = 26$ ) had a lower general QOL score and physical health than those who reported one ADE ( $N = 11$ ) (median of 6.5 versus 8.0 on the general QOL score,  $P = 0.032$ ; mean of 13.6 versus 14.9 on physical health,  $P = 0.138$ ).

**Table 2.1.** Patient characteristics and nature of the reported ADEs

Characteristic	N (%)
Patients	135
Women	47 (35)
Mean age in years (SD)	64 (9)
Education	
Low <sup>a</sup>	34 (25)
Middle <sup>b</sup>	53 (39)
High <sup>c</sup>	39 (29)
Other	9 (7)
Number of prescribed oral glucose-lowering drugs <sup>d</sup>	
1 oral glucose-lowering drug	91 (68)
2 oral glucose-lowering drugs	38 (29)
3 oral glucose-lowering drugs	4 (3)
Classes of prescribed oral glucose-lowering drugs <sup>d</sup>	
Biguanides	125 (94)
Dipeptidyl peptidase 4 (DPP-4) inhibitors	4 (3)
Glucagon-like peptide-1 (GLP-1)	5 (4)
Sulfonamides	42 (32)
Thiazolidinediones	3 (2)
Median number of systemic drugs prescribed (IQR)	7 (5–9)
Patients additionally on insulin	24 (18)
Patients reporting an ADE	37 (27)
Number of ADEs reported (range)	146 (1–11)
First time experiencing the ADE	
Today	6 (4)
Yesterday	1 (1)
2–7 days ago	5 (3)
Between 1 week and 1 month ago	14 (10)
Between 1 and 6 months ago	14 (10)
Between 6 and 12 months ago	21 (14)
More than 12 months ago	85 (58)
ADE gone away or improved	
Not yet	120 (82)
Clearly improved	14 (10)
ADE was treated and has improved	4 (3)
ADE went away by itself	0 (0)
ADE went away after quitting medication	1 (1)
ADE went away after treatment	1 (1)
Other	6 (4)
How much bothersome	
Not at all	11 (8)
Only a bit	22 (15)
Somewhat	71 (49)
Quite a lot	31 (21)
Very much	11 (8)
Influence daily functioning	
None	47 (32)
Only a bit	24 (17)
Somewhat	54 (37)
Quite a lot	18 (12)
Very much	2 (1)

ADEs = Adverse drug events; SD = Standard deviation; IQR = Interquartile range

<sup>a</sup> No education; elementary school; junior secondary vocational education

<sup>b</sup> Junior general secondary education; senior secondary vocational education

<sup>c</sup> Senior general secondary education; higher professional education; university education

<sup>d</sup> Medication use is based on 133 patients since drug information of 2 patients was not available

**Table 2.2.** QOL and differences between patients who report and those who do not report an ADE

	Total	Patients not reporting an ADE	Patients reporting $\geq 1$ ADEs	P-value	Effect size <i>d</i>
Number of patients	135	98	37		
Mean QOL (SD)					
Median general QOL score (IQR)	8.0 (6.0-8.0)	8.0 (7.0-8.0)	7.0 (6.0-8.0)	<b>0.021<sup>†</sup></b>	-0.53
Physical health	14.9 (2.5)	15.2 (2.4)	14.0 (2.6)	<b>0.009*</b>	-0.50
Psychological health	14.7 (2.4)	15.0 (2.4)	14.1 (2.5)	0.064*	-0.38
Social relationships	14.5 (2.6)	14.6 (2.7)	14.4 (2.6)	0.666*	-0.07
Environment	16.0 (2.2)	16.2 (2.2)	15.6 (2.3)	0.134*	-0.27

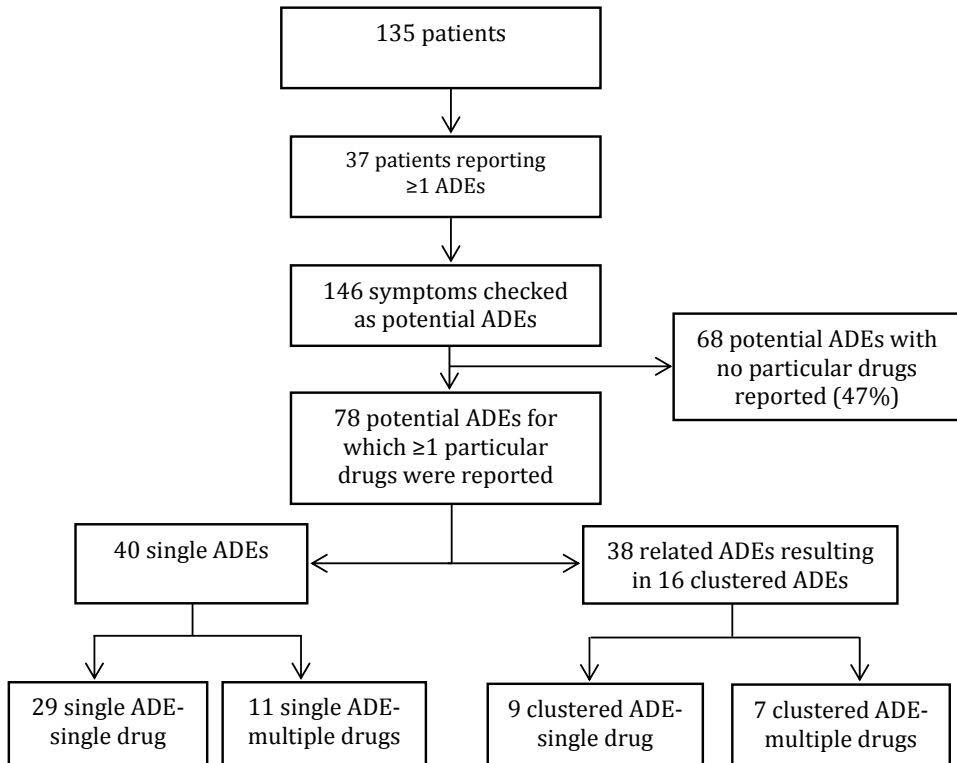
QOL = Quality of life; ADE = Adverse drug event; SD = Standard deviation; IQR = Interquartile range.  
<sup>†</sup> Mann-Whitney U test; \* T-test.

### Concurrent validity

#### *Comparison of patient-reported ADE-drug association with known ADEs*

For 78 reported ADEs (53%), patients mentioned at least one particular drug for the ADE. These patients (N = 25) did not significantly differ in age, gender, and education level from the patients who did not mention a particular drug for any of their ADEs (N = 12) (**Appendix 3**; supplemental table 2). For the ADEs that were associated with one or more particular drugs (N = 78), it was more often reported that the ADE occurred (soon) after the start or dosage increase of a drug and that a healthcare professional confirmed the symptom being an ADE than for the ADEs not associated with a particular drug (N = 68) (**Appendix 3**; Supplemental table 3). ADEs that were already perceived before the start of a drug but increased in frequency after the start of a drug were less likely to be associated with a particular drug. For 15% of both ADEs, those that were associated with a particular drug and those that were not, the patients reported that the symptom was described in the patient leaflet as being an ADE for that particular drug.

For almost half of the ADEs that were associated with one or more particular drugs (38 cases; 49%), multiple related ADEs describing one overall ADE were observed resulting in 16 clustered ADE-drug associations (**Figure 2.1**). All of the single and clustered ADE-drug associations (N = 56) were compared with the known ADEs for those drugs as documented in the SPC (**Appendix 3**; supplemental table 4). In 41 of the cases (73%), the ADE-drug association was in agreement with known ADEs for at least one of the reported drugs, indicating sufficient concurrent validity.



**Figure 2.1.** Flow-chart of reported adverse drug events (ADEs) and ADE-drug associations

There were 29 single ADE-single drug associations. Of these ADEs, 22 (76%) were in agreement with the known ADEs. The patient-reported causality score for these single ADE-single drug associations ranged from 1 to 3 with a median of 2 (interquartile range: 1.0 – 2.5). The single ADE-single drug associations with a causality assessment score  $\geq 2$  ( $N = 21$ ) were in 76% of the cases ( $N = 16$ ) in agreement with the known ADEs. This agreement increased to 100% when only causality scores of 3 ( $N = 7$ ) were taken into account.

*Comparison of ADE-metformin association between ADE questionnaire and treatment/drug-specific questionnaire*

Of the 135 patients, 125 reported that they had used metformin in the previous 4 weeks. The comparison of reporting an ADE with metformin use between the ADE questionnaire and the TSQM revealed 11 true positives, 93 true negatives, 3 false positives, and 18 false negatives. These numbers resulted in a sufficient, high positive predictive value (79%; 95% confidence interval 49-95%), but an insufficient, low sensitivity (38%; 95% confidence interval 21-58%).

## Discussion

Construct validity of the ADE questionnaire was demonstrated by an association between ADEs and QOL, where patients who reported an ADE on the questionnaire had a lower general QOL and physical health than patients who did not report an ADE. The 73% agreement between the ADE-drug associations as reported by the patients and the known ADEs for those drugs as documented in the SPC indicates concurrent validity. However, additional concurrent validity assessment revealed that the ADE questionnaire has sufficient positive predictive value but insufficient sensitivity when comparing the linkage of an ADE to metformin in the ADE questionnaire with the treatment/drug-specific TSQM questionnaire.

For the construct validity, patients who reported an ADE had statistically significant lower scores than those not reporting an ADE on physical health and on the general QOL score. The lowest QOL scores were observed for patients reporting more than 1 ADE. The effect sizes indicate that the QOL differences were clinically relevant. The underlying mechanism for the association between ADEs and QOL is not fully understood yet. Although, an increase in the experience of ADEs may be associated with a decrease in QOL <sup>[11,12]</sup>, it has also been shown that baseline QOL scores were lower for patients who report an ADE at follow-up than for patients who do not report an ADE at follow-up <sup>[34]</sup>. This finding may imply that patients with a low QOL have a higher risk of experiencing ADEs <sup>[35]</sup>, or that patients with negative health perceptions or certain personality traits are more disposed to experiencing and/or reporting ADEs <sup>[13,36]</sup>. The underlying mechanism for the association between ADEs and QOL does not influence its usefulness to assess construct validity, since the principle of construct validity is based on the existence of an adequate association between scores on instruments <sup>[37]</sup>.

Agreement of at least one drug that patients mentioned for their ADE on the patient-reported ADE questionnaire with known ADEs as documented in the SPC supports its concurrent validity. Full agreement was shown when only cases with a high patient-reported causality score were included, indicating that a patient-reported causality assessment increases the validity of ADE reporting by patients. Previously, it was shown that patients cannot perceive all ADEs and also do not always make the connection between a symptom and their drug use <sup>[7]</sup>. In addition, we found that only 53% of the patients who reported an ADE mentioned a particular drug for the ADE. This low number may especially occur in a patient population, such as included in this study. Patients with diabetes using oral glucose-regulating drugs often use multiple drugs for various related and unrelated diseases, which complicates the assignment of a drug to the symptom <sup>[38]</sup>. We found that mentioning a particular drug was more likely when a healthcare professional confirmed the symptom being an ADE. This finding indicates that patient-reported ADE instruments cannot fully replace healthcare professional

reports. Special attention of healthcare professionals may be necessary for symptoms that are already present but increase in frequency after starting a drug, and for ADEs that do not occur soon after the start of a drug. Our study results support that for reliable knowledge about ADEs, information from both healthcare professionals and patients is needed <sup>[39]</sup>.

For 15% of the ADEs, the patients reported that the ADE was described in the patient leaflet. Surprisingly, this was also observed for ADEs for which patients did not mention a particular drug. It may be that the ADE is reported in the patient leaflet of multiple drugs and the patient is unsure about the exact drug that may cause the ADE. Additional preferably qualitative research is needed to better understand these reports.

Comparing the ADE questionnaire with the TSQM only partly supported the concurrent validity. The positive predictive value indicates that an ADE with the use of metformin detected by the ADE questionnaire is likely to be an ADE. The low sensitivity, on the other hand, implies that not all experienced ADEs associated with metformin use are detected by the ADE questionnaire. There are, however, notable differences between the ADE questionnaire and the TSQM that may negatively influence the validity measures. First, the patient-reported ADE questionnaire includes a recall period of 4 weeks whereas the TSQM does not include a recall period. Second, in the patient-reported ADE questionnaire it is asked which drug is associated with the ADE, whereas in the TSQM it is asked whether or not an ADE is experienced with the use of the metformin.

More than half of the ADEs reported in this study started more than 12 months ago. This timing of the ADE may have influenced our validity assessment. For construct validity, the experience of an ADE for a long period of time may lead to a stronger association with QOL due to the longer period of burden but could also lead to a weaker association due to adaptation of getting used to the symptoms. The timing of the ADE seemed to influence whether or not a particular drug was mentioned for the ADE (**Appendix 3**; supplemental table 3). This finding indicates that particularly the more recent ADEs were included in the concurrent validity assessment, since only patients who report a specific drug were included in this assessment. It is not clear whether this could have impact on the associations with the SPC and the TSQM conducted for the concurrent validity.

An important limitation of this study is the low number of patients reporting an ADE. The construct validity of the ADE questionnaire was only assessed by testing the association between patient-reported ADEs and QOL. Future studies should assess the construct validity using additional constructs, also at specific ADE level and taking the severity of the ADE into account. The severity of an ADE has shown to influence a patient's QOL <sup>[35]</sup>. Participating patients in this study were significantly younger than

the non-responders <sup>[17]</sup>, and were more often prescribed metformin and less often a sulfonamide or thiazolidinedione than other, on average elder, patients with diabetes in the Dutch primary care <sup>[40]</sup>. These differences may be due to the use of a web-based version of the questionnaire, because it has been shown that respondents to web-based questionnaires are younger and have a shorter diabetes duration <sup>[41,42]</sup>. It should also be noted that the validity of the paper-based version of the questionnaire may differ from the web-based version due to differences between these methods in, for instance, branching of questions, visual aspects, or respondent required actions such as the use of a mouse in the web-based version <sup>[43-45]</sup>. Another limitation was that about half of the reported ADEs that were associated by the patients with particular drugs were assessed by the researchers as being related ADEs describing one overall ADE. In this study, these ADEs were clustered by the researchers. We recommend that future research with checklist-based patient-reported ADE questionnaires should include the option for patients to cluster related ADEs. This option will likely increase the validity of the instrument and the practicability for the researchers. We included patients using drugs for type 2 diabetes in this study, who particularly reported ADEs belonging to the gastrointestinal disorders SOC of the MedDRA<sup>®</sup> <sup>[17]</sup>. Although these patients also often used drugs for other diseases, the use of the questionnaire in other patient populations and in the general population in which other ADEs are also common requires additional validation <sup>[10]</sup>. The validity of the questionnaire is expected to differ especially for more serious ADEs (e.g. rhabdomyolysis with statin use) and for ADEs that can be distinguished more clearly from the symptoms related to the disease (e.g. alopecia with the use of chemotherapy in patients with cancer). Finally, it should be noted that this validation was conducted in an observational study, where most patients were current users and familiar with their drug treatment. Therefore, additional studies are needed assessing the validity of the instrument in a clinical trial, where patients are often new users and blinded to the drug treatment. However, the insufficient concurrent validity indicates that the first next step is to make adaptations to the questionnaire.

## **Conclusions**

Our results indicate that the patient-reported ADE questionnaire has construct validity of reporting any ADE versus no ADE, since the reporting of an ADE was associated with a lower QOL. Additional assessment is needed to test the construct validity for individual ADEs. The concurrent validity was only partly demonstrated by 1) sufficient agreement between the ADE-drug associations and the known ADEs when the researchers clustered related ADEs into one ADE, 2) sufficient positive predictive value to detect 'real' ADEs associated with metformin use, but 3) insufficient sensitivity to detect all ADEs that are associated with metformin use. Therefore, adaptations to the ADE questionnaire, such

as allowing patients to cluster multiple related ADEs as one overall ADE, need to be made and tested in future validation studies.

### **Acknowledgements**

MedDRA® is a registered trademark of the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA). We would like to thank the WHO for providing the WHOQOL-BREF questionnaire and Quintiles for providing the TSQM.



## References

- [1] Suvarna V. Phase IV of drug development. *Perspect Clin Res* 2010;1(2):57–60.
- [2] Berlin JA, Glasser SC, Ellenberg SS. Adverse event detection in drug development: recommendations and obligations beyond phase 3. *Am J Public Health* 2008;98(8):1366–71.
- [3] Walshe K. Adverse events in health care: issues in measurement. *Qual Health Care* 2000;9(1):47–52.
- [4] Basch E, Iasonos A, McDonough T, Barz A, Culkun A, Kris MG, et al. Patient versus clinician symptom reporting using the national cancer institute common terminology criteria for adverse events: results of a questionnaire-based study. *Lancet Oncol* 2006;7(11):903–9.
- [5] Blenkinsopp A, Wilkie P, Wang M, Routledge PA. Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. *Br J Clin Pharmacol* 2007;63(2):148–56.
- [6] Hakobyan L, Haaijer-Ruskamp FM, de Zeeuw D, Dobre D, Denig P. Comparing adverse event rates of oral blood glucose-lowering drugs reported by patients and healthcare providers: a post-hoc analysis of observational studies published between 1999 and 2011. *Drug Saf* 2011;34(12):1191–202.
- [7] Gawert L, Hiersche F, Zink A, Strangfeld A. How well do patient reports reflect adverse drug reactions reported by rheumatologists? Agreement of physician- and patient-reported adverse events in patients with rheumatoid arthritis observed in the German biologics register. *Rheumatology (Oxford)* 2011;50(1):152–60.
- [8] US Department of Health and Human Services, FDA, Center for Drug Evaluation and Research Center for Biologics Evaluation and Research Center for Devices and Radiological Health. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>.
- [9] European Medicines Agency. Fourth report on the progress of the interaction with patients' and consumers' organizations (2010) and results/analysis of the degree of satisfaction of patients and consumers involved in EMA activities during 2010. EMA/632696/2011. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Report/2011/10/WC500116866.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Report/2011/10/WC500116866.pdf).
- [10] Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* 2007;10(Suppl 2):94–105.
- [11] Azevedo AM, Alonso NB, Vidal-Dourado M, Noffs MH, Pascalicchio TF, Caboclo LO, et al. Validity and reliability of the Portuguese-Brazilian version of the quality of life in epilepsy inventory-89. *Epilepsy Behav* 2009;14(3):465–71.
- [12] Carreno M, Donaire A, Falip M, Maestro I, Fernandez S, Nagel AG, et al. Validation of the Spanish version of the Liverpool adverse events profile in patients with epilepsy. *Epilepsy Behav* 2009;15(2):154–9.
- [13] De Smedt RH, Denig P, Haaijer-Ruskamp FM, Jaarsma T. Perceived medication adverse effects and coping strategies reported by chronic heart failure patients. *Int J Clin Pract* 2009;63(2):233–42.
- [14] Foster JM, van Sonderen E, Lee AJ, Sanderman R, Dijkstra A, Postma DS, et al. A self-rating scale for patient-perceived side effects of inhaled corticosteroids. *Respir Res* 2006;7:131.
- [15] Jarernsiripornkul N, Krska J, Capps PA, Richards RM, Lee A. Patient reporting of potential adverse drug reactions: a methodological study. *Br J Clin Pharmacol* 2002;53(3):318–25.
- [16] Corso DM, Pucino F, DeLeo JM, Calis KA, Gallelli JF. Development of a questionnaire for detecting potential adverse drug reactions. *Ann Pharmacother* 1992;26(7-8):890–6.
- [17] De Vries ST, Mol PG, de Zeeuw D, Haaijer-Ruskamp FM, Denig P. Development and initial validation of a patient-reported adverse drug event questionnaire. *Drug Saf* 2013;36:765–77.
- [18] Unipark Enterprise Feedback Suite 8.0 version 1.1. <http://www.unipark.de>.
- [19] De Vries J, Van Heck GL. WHOQOL-Bref. Tilburg: Universiteit van Tilburg; 1996.

- [20] Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The world health organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Qual Life Res* 2004;13(2):299-310.
- [21] Dutch Medicines Evaluation Board. <http://www.cbg-meb.nl/cbg/nl>.
- [22] Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981;30(2):239-45.
- [23] Krska J, Anderson C, Murphy E, Avery AJ. How patient reporters identify adverse drug reactions: a qualitative study of reporting via the UK yellow card scheme. *Drug Saf* 2011;34(5):429-36.
- [24] Atkinson MJ, Sinha A, Hass SL, Colman SS, Kumar RN, Brod M, et al. Validation of a general measure of treatment satisfaction, the treatment satisfaction questionnaire for medication (TSQM), using a national panel study of chronic disease. *Health Qual Life Outcomes* 2004;2:12.
- [25] American Diabetes Association: Standards of medical care in diabetes-2013. *Diabetes Care* 2013;36(Suppl 1):11-66.
- [26] Rutten GEHM, De Grauw WJC, Nijpels G, Houweling ST, Van de Laar FA, Bilo HJ, et al. The NHG guideline diabetes mellitus type 2. *Huisarts Wet* 2013;56:512-25.
- [27] Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T, Clinical Significance Consensus Meeting Group. Estimating clinically significant differences in quality of life outcomes. *Qual Life Res* 2005;14(2):285-95.
- [28] Aursnes I, Gjertsen MK. Common adverse events associated with an SSRI: meta-analysis of early paroxetine data. *Pharmacoepidemiol Drug Saf* 2008;17(7):707-13.
- [29] Bergk V, Haefeli WE, Gasse C, Brenner H, Martin-Facklam M. Information deficits in the summary of product characteristics preclude an optimal management of drug interactions: a comparison with evidence from the literature. *Eur J Clin Pharmacol* 2005;61(5-6):327-35.
- [30] Aagaard L, Hansen EH. Consumers' reports of suspected adverse drug reactions volunteered to a consumer magazine. *Br J Clin Pharmacol* 2010;69(3):317-8.
- [31] Altman DG, Bland JM. Diagnostic tests: 1: sensitivity and specificity. *BMJ* 1994;308(6943):1552.
- [32] Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ* 1994;309(6947):102.
- [33] Deeks JJ, Altman DG. Sensitivity and specificity and their confidence intervals cannot exceed 100%. *BMJ* 1999;318(7177):193-4.
- [34] Ernst EJ, Ernst ME, Hoehns JD, Bergus GR. Women's quality of life is decreased by acute cystitis and antibiotic adverse effects associated with treatment. *Health Qual Life Outcomes* 2005;3:45.
- [35] Guo N, Marra F, Fitzgerald JM, Elwood RK, Marra CA. Impact of adverse drug reaction and predictivity of quality of life status in tuberculosis. *Eur Respir J* 2010;36(1):206-8.
- [36] Imayama I, Plotnikoff RC, Courneya KS, Johnson JA. Determinants of quality of life in type 2 diabetes population: the inclusion of personality. *Qual Life Res* 2011;20(4):551-8.
- [37] De Vet H, Terwee C, Mokkink L, Knol D. *Measurement in Medicine: A practical guide*. Cambridge: University Press; 2011.
- [38] Chaipichit N, Krska J, Pratipanawatr T, Uchaipichat V, Jarernsiripornkul N. A qualitative study to explore how patients identify and assess symptoms as adverse drug reactions. *Eur J Clin Pharmacol* 2014;70(5):607-15.
- [39] Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *J Natl Cancer Inst* 2009;101(23):1624-32.
- [40] Voorham J, Haaijer-Ruskamp FM, Wolffenbuttel BH, Stolk RP, Denig P, Groningen Initiative to Analyze Type 2 Diabetes Treatment Group. Medication adherence affects treatment modifications in patients with type 2 diabetes. *Clin Ther* 2011;33(1):121-34.
- [41] Härmark L, Alberts S, van Puijenbroek E, Denig P, van Grootheest K. Representativeness of diabetes patients participating in a web-based adverse drug reaction monitoring system. *Pharmacoepidemiol Drug Saf* 2013;22(3):250-5.
- [42] Schiøtz M, Bøgelund M, Willaing I. Challenges using online surveys in a Danish population of people with type 2 diabetes. *Chronic Illn* 2012;8(1):56-63.

- [43] Wyatt JC. When to use web-based surveys. *J Am Med Inform Assoc* 2000;7(4):426–9.
- [44] Manfreda KL, Batagelj Z, Vehovar V. Design of Web Survey Questionnaires: Three Basic Experiments. *J Computer-Mediated Commun* 2002;7.
- [45] Couper MP. Web surveys: a review of issues and approaches. *Public Opin Q* 2000, 64(4):464–494.