The Global Usability Score Short-Form for the simplified assessment of dry powder inhalers (DPIs) usability

Massimiliano Povero,1 Paola Turco,2 Luca Bonadiman,1 Roberto W. Dal Negro3
1AdRes Health Economics and Outcome Research, Turin
2Research & Clinical Governance, Verona
3National Centre for Respiratory Pharmacoeconomics and Pharmacoepidemiology, Verona, Italy

Background: The choice of the Dry Powder Inhaler (DPI) to prescribe is a critical issue. The estimation of DPIs usability depends on the objective assessment of several indices related to both subjective and objective determinants. The Global Usability Score (GUS) Questionnaire is a comprehensive tool usable for checking, comparing, and ranking inhalers’ usability objectively in real life, but it takes some time to fill.

Aim: The aim of this study was to favour the quicker check of DPIs usability in clinical practice by means of a simplified short-form GUS (S-GUS) Questionnaire, while maintaining the high specificity and sensitivity of the original, extended version of the Questionnaire (O-GUS questionnaire).

Methods: The usability of the six most prescribed DPIs was assessed in 222 patients with persistent airway obstruction and needing long-term inhalation treatments. LASSO regression and multicollinearity test were used to select the subset of questions of the O-GUS questionnaire, with the highest information power. Each item was then scored using the corresponding coefficient in the linear regression (normalized at 50 as the O-GUS score). Agreement between the original and the short-form questionnaire was evaluated using the Cohen’s kappa statistic (κ). The overall S-GUS values obtained for each DPI were then compared to those from the O-GUS, in the same patients, using a Bayesian indirect comparison (IC) model.

Results: After the statistical selection of the items mostly contributing to the overall score, the novel S-GUS questionnaire consists of twelve items only. Nine items are related to patients’ opinion before DPIs handling, and three to the nurse’s assessment after DPIs practicality. O-GUS and S-GUS score were strongly correlated (R²=0.9843, p<0.0001) and the usability score calculated for each DPI by means of the O- and of S-GUS overlapped almost completely (κ=84.5%, 95% CI 81.3% to 89.2%). Furthermore, S-GUS was much faster to complete than O-GUS (mean time 6.1 vs 23.4 minutes, p<0.001). Estimates of S-GUS, obtained from the IC model, allowed to propose a simple classification of usability: “good” by GUS values >25; “pretty good” by values ≤25≥15, and “insufficient” by values <15.

Conclusions: The S-GUS proves as much specific and suitable as the extended O-GUS questionnaire in measuring DPIs usability, while maintaining the same high sensitivity. As the time required for its use is quite shorter, S-GUS is also particularly suitable and helpful in current clinical practice.

Key words: Inhalation devices; dry powder inhalers; usability; global usability score; GUS; GUS short-form; bronchial asthma; COPD.

Correspondence: Massimiliano Povero, AdRes Health Economics and Outcome Resources, Via Vittorio Alfieri 17, 10121 Torino, Italy. E-mail: m.povero@adreshe.com

Contributions: All authors read and approved the final manuscript.

Conflict of interest: The authors declare that they have no competing interests. RWD is Associate Editor of Multidisciplinary Respiratory Medicine.

Funding: Not applicable.

Availability of data and materials: Authors do not wish to share their data without their permission.

Ethics approval and consent to participate: This study was approved by the Ethical and Scientific Commission of the National Centre for Respiratory Pharmacoeconomics and Pharmacoepidemiology during the session of January 4th, 2016 (# RD 011/G01/16). All subjects gave their written informed consent.

Consent for publication: All subjects gave their written informed consent to the study.
Introduction

The evidence that inhalers, and Dry Powder Inhalers (DPIs) in particular, may affect per sé therapeutic outcomes independently of the drug(s) prescribed is consolidated [1-5]. On the other hand, it has also been stated that patients are not able to use all DPIs equally well [5-8]. As a consequence, the achievement of a deeper awareness of inhalers’ usability should then be increasingly pursued together with a dedicated specific training as priority issues [6-9]. The empowerment of patients who need inhalation therapy has risen remarkably in the last decade, particularly in favour of those individuals requiring long-term treatments for managing their persistent airway obstruction (such as: Bronchial Asthma and Chronic Obstructive Pulmonary Disease - COPD).

The choice of the convenient DPI to prescribe still represents a crucial issue in real-life because it is too frequently affected by highly empiric decisions in clinical practice [7-8].

The critical factors contributing to the effectiveness of inhalation treatments via DPIs are various and of different origin (Figure 1). Differently from the past, the primary role of patients’ viewpoint greatly concentrated the attention in recent years, and the determinants of incorrect inhalation were frequently investigated [9-10]. Patients’ preference, acceptance, and satisfaction were mostly assigned to criteria strictly related to their current subjectivity and beliefs (such as: age, psychological profile, socio-economic status, cognition, educational level, physical limitations) [11-17], despite the low number of studies assessing the objective correspondence between the patients’ beliefs and the real-life usability of inhalers by means of quantitative tools [11,14,18]. However, unlike those belonging to the domain of subjectivity, other criteria are involved in determining the appropriate utilization of DPIs, particularly those that are almost completely independent of the sole patients’ point of view: the intrinsic structure of the device, the number and the complexity of manoeuvres required for their actuation, the quality and the duration of the educational training received, etc.

Figure 1. The multiple aspects affecting DPIs usability.

Novel and more comprehensive instruments of investigation should then be suggested and expected in order to provide a more objective check of inhalers usability. At present, the Global Usability Score (GUS), which is obtained by means of the specific GUS Questionnaire, represents the first comprehensive tool for assessing, ranking, and comparing inhalers’ usability objectively, thus greatly contributing to the effective and motivated standard of choice in real life [19-23]. However, even if characterized by a >97% comprehension at the first reading, the original, extended version of GUS Questionnaire (O-GUS) consisting of 27 questions, obviously requires some time to be properly filled, and this fact may likely limit its wider use in clinical practice.

Aim of the present study was the definition and the validation of the simplified short-form GUS Questionnaire (S-GUS) in order to favour the effective and quick assessment of DPIs usability in daily clinical activities, while maintaining the same high specificity and sensitivity of the O-GUS Questionnaire.

Materials and Methods

A cluster of 222 original, extended GUS questionnaires (consisting in 27 items; 22 questions requiring the patients’, and the remaining 5 the nurse’s direct involvement) already used in a previous investigation [24] aimed to assess and compare the usability of the six most used DPIs (Breezhaler, Diskus, Ellipta, Nexthaler, Spiromax, and Turbuhaler randomly grouped) was used. This study was approved by the Ethical and Scientific Commission of the National Centre for Respiratory Pharmacoeconomics and Pharmacoepidemiology during the session of January 4th, 2016 (# RD 011/G01/16).

The association between the 27 items of the O-GUS questionnaire and the GUS scores was investigated using a Least Absolute Shrinkage and Selection Operator (LASSO) regression model. The LASSO regression minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant [25]. Because of the nature of this constraint, it tends to produce some coefficients that are exactly zero and hence it permits to discard redundant covariates. LASSO regression is recommended with respect to automated variable selection algorithms as backward elimination, forward selection or stepwise
because this approach avoids multiple test bias [26,27]. Multicollinearity was also tested using the Variance Inflation Factors (VIF) and covariates with VIF>5 were discarded [28]; the procedure corresponds to exclude the questions which proved characterized by an equivalent information value. Finally, the S-GUS questionnaire was constructed using such subset of covariates, and coefficients of linear regression were used for scoring each item (such that the score ranges between 0 and 50 as the previous one). The responses of the 222 O-GUS questionnaire were used to calculate the new S-GUS questionnaire and original vs short-form scores were compared using univariate linear regression in order to test the correlation between the two questionnaires. Agreement between O-GUS and S-GUS was defined as the ability of ranking in the same way two different devices i.e. if O-GUS of device A is greater than O-GUS of device B then S-GUS of device A should be greater than S-GUS of device B. To take into account the agreement occurring by chance we evaluated the concordance between original and short-form questionnaire using the Cohen’s kappa statistics (κ). Because of the high number of possible comparisons (222 questionnaires + 4 devices for each questionnaire = 393,828 comparisons) we preferred to calculate confidence interval for κ using bootstrapping approach with 1,000 replications, sampling for each replication a subset of 25 questionnaires corresponding to 100 tested devices (4 for each questionnaire) and 4,950 comparisons.

Furthermore, all pairwise comparisons between the six DPIs (using both questionnaires) were merged using a Bayesian Indirect Comparison (IC) model; details of the model were previously described elsewhere [24,29]. Briefly, IC models are used for pooling quantitative results from multiple studies and for assessing the effect between two or among more treatments; in our context, “multiple studies” means that each patient tested more than two devices, and “treatments” means devices. This approach is particularly advantageous because all devices under comparison are incorporated into a single model even if they are not compared in the same questionnaire. Furthermore, to rank the treatments, we used the Surface Under the Cumulative RAnking probabilities (SUCRA), the larger the SUCRA, the higher the device global usability [30]. Both fixed-effect (FE) and random-effect (RE) models were performed, and the model with the lowest Deviance Information Criterion (DIC) was selected [31]. DIC represents a trade-off between goodness of fit (how well the model fits the observations) and model complexity (number of model parameters).

Finally, the time for completing both O-GUS and S-GUS was also counted in minutes, and compared by two-side t-test. Furthermore, the results of IC model constructed on the S-GUS scores were used to propose a simply classification of “Good”, “Pretty good” and “Insufficient” usability for each tested device.

Results

The items characterized by the highest informative power and then the most suitable to contribute to the novel S-GUS Questionnaire composition, are reported in Table 1 together with the corresponding linear coefficients. Figure 2 illustrates the resulted S-GUS questionnaire with the new scores for each question.

GUS Questionnaire short-form

Before DPIs utilization

1. Which device do you prefer "at glance"?
   A B C D 1.4
2. Which device do you perceive as the easiest to use?
   A B C D 3.5
3. Which device do you perceive as the most difficult to use?
   A B C D 2.5*

Which device do you prefer (only one exclusive choice/item) in terms of:

4. Shape
   A B C D 2.4
5. Mouthpiece
   A B C D 1.4
6. Hygiene
   A B C D 3
7. Presence of dose counter
   A B C D 2.9
8. Ease of gripping
   A B C D 4.6
9. Number of maneuvers for actuation
   A B C D 12.8

After DPIs utilization

10. Number of attempts for the 1st proper actuation with each device
    7.3 4.3 1.7 0
    1 2 3 >3

11. Time (in minutes) for the 1st proper actuation with each device
    5 2.6 1.4 0
    <2 <3 <6 >6

12. Concordance between patient and nurse on the easiest DPI
    3.2 0
    Y N

Figure 2. GUS Questionnaire short-form with only questions characterized by the highest degree of linear association (in bold the score for each question). *+2.5 points for each device not perceived as the most difficult to use.

The items of the new questionnaire were reduced to twelve, such as only those highly contributing to the final score, and then unavoidable: nine were dealing to the patients’ opinion before DPIs handling, and three after DPIs utilization. While the former items were uniquely related to patients’ perceptions and beliefs, the latter three were linked to the quantitative assessment by the expert nurse, carefully supervising all inhalation procedures objectively.

A single sub-score each scored the patients’ answers to the first nine items. Items #10 and #11 were scored by decreasing values according to the degree of difficulties objectively encountered by patients with each DPI, while a categorical sub-score was assigned to item #12, such as 3.2 (Yes) or 0 (No), in case of agreement or not agreement between patient’s and nurse’s judgment, respectively (Figure 2). The overall S-GUS summed up to 50 points as in the case of the extended O-GUS.

Linear regression analysis between O-GUS and S-GUS scores proved a very strong correlation (R²=0.9843, p<0.0001) and the ratio between S-GUS and O-GUS score was almost 1 (0.9904, p<0.0001) (Figure 3). Indeed, the crude concordance between original and short-form is 93.3% (95% CI 91.6% to 95.2%), and such agreement remains very high even after removing the agreement occurring by chance (κ=84.5%, 95% CI 81.3% to 89.2%).

Figure 3. Linear regression calculated between the original and the simplified GUS in all 222 subjects (888 measures, 4 devices for each subjects).

Table 1. Reduced linear model for the determination of the short-form GUS questionnaire.

<table>
<thead>
<tr>
<th>S-GUS attributes for each device</th>
<th>Estimate (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred “at glance” (yes vs no)</td>
<td>1.43 (0.68 to 2.18)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Perceived as the easiest to use (yes vs no)</td>
<td>3.69 (2.82 to 4.58)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Perceived as the most difficult to use (yes vs no)</td>
<td>2.58 (1.94 to 3.23)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Preferred in terms of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape (yes vs no)</td>
<td>2.57 (2.02 to 3.12)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mouthpiece (yes vs no)</td>
<td>3.16 (2.44 to 3.87)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hygiene (yes vs no)</td>
<td>3.1 (2.32 to 3.85)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dose counter (yes vs no)</td>
<td>4.81 (4.15 to 5.47)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ease of gripping (yes vs no)</td>
<td>13.43 (12.64 to 14.22)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Number of attempts for the 1st proper actuation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One (≥ 3 more than three)</td>
<td>7.73 (4.03 to 11.43)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Two (≥ 3 more than three)</td>
<td>4.5 (0.8 to 8.21)</td>
<td>0.0172</td>
</tr>
<tr>
<td>Three (≥ 3 more than three)</td>
<td>1.83 (-1.99 to 5.65)</td>
<td>0.3473</td>
</tr>
<tr>
<td>Time for the 1st proper actuation*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 2 min (vs more than 6 min)</td>
<td>5.22 (3.99 to 6.44)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>2-3 min (vs more than 6 min)</td>
<td>2.75 (1.87 to 3.62)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>3-6 min (vs more than 6 min)</td>
<td>1.47 (0.95 to 1.99)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Concordance nurse-patient on the easiest DPI</td>
<td>3.42 (2.62 to 4.22)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

CI, confidence interval; DPI, dry powder inhaler; * defined as the total time from the beginning of the first nurse’s demonstration to the correct actuation (nurse’s demonstration was repeated in case of wrong actuation).

Table 2. Estimated GUS and ranking of competing devices using both original and short-form GUS questionnaire.

<table>
<thead>
<tr>
<th>Device</th>
<th>Median GUS (95% CrI)</th>
<th>O-GUS SUCRA</th>
<th>Median rank (95% CrI)</th>
<th>Median GUS (95% CrI)</th>
<th>S-GUS SUCRA</th>
<th>Median rank (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellipta</td>
<td>26.9 (20 to 33)</td>
<td>91.4%</td>
<td>1 (1 to 3)</td>
<td>27.4 (21 to 34)</td>
<td>83.3%</td>
<td>1 (1 to 3)</td>
</tr>
<tr>
<td>Turbhaler</td>
<td>23.9 (17 to 31)</td>
<td>74.8%</td>
<td>2 (1 to 4)</td>
<td>23.7 (17 to 31)</td>
<td>74.4%</td>
<td>2 (1 to 4)</td>
</tr>
<tr>
<td>Diskus</td>
<td>23.4 (17 to 30)</td>
<td>72%</td>
<td>2 (1 to 4)</td>
<td>22.9 (16 to 30)</td>
<td>68.8%</td>
<td>3 (1 to 4)</td>
</tr>
<tr>
<td>Spiromax</td>
<td>17 (12 to 22)</td>
<td>37.4%</td>
<td>4 (3 to 5)</td>
<td>18.2 (13 to 23)</td>
<td>39.4%</td>
<td>4 (3 to 5)</td>
</tr>
<tr>
<td>Nexthaler</td>
<td>14.5 (8 to 21)</td>
<td>24.3%</td>
<td>5 (4 to 5)</td>
<td>15.4 (9 to 22)</td>
<td>24.1%</td>
<td>5 (3 to 5)</td>
</tr>
<tr>
<td>Breezhaler</td>
<td>5.2 (5 to 6)</td>
<td>0.1%</td>
<td>6 (6 to 6)</td>
<td>6.4 (6 to 7)</td>
<td>0.1%</td>
<td>6 (6 to 6)</td>
</tr>
</tbody>
</table>

CrI, credibility interval; GUS, global usability score; O-GUS, original global usability score; S-GUS, short-form global usability score; SUCRA, surface under the cumulative ranking curves.
Table 2 displays the results produced by pairwise IC both using O-GUS and S-GUS scores. There is absolute perfect agreement between the results obtained using O-GUS score and those obtained using S-GUS score. Furthermore, devices are ranked in the same order, the value of SUCRA for each DPI is almost equal and the cumulative ranking probability curves overlap for each DPI (Figure 4). Mean values of the overall S-GUS proved stable for all DPIs and characterized by a well comparable variability. This stability allows the suggestion of a quantitative ranking of DPIs based on the mean S-GUS value. In particular, a “good” usability can be highly presumed for S-GUS values ≥25, and a “pretty good” usability can be likely suggested for S-GUS values <25 (Figure 5)

Finally, the mean time spent for filling O-GUS properly was 23.4 min (standard deviation 3.3), while that for completing the S-GUS was 6.1 min (standard deviation 2.2), respectively (p<0.001).

Figure 4. Cumulative ranking probability plots for competing DPIs drawn by extrapolating at the middle of each interval of the step function: on the horizontal axis is the possible rank of each device (from the first to the worst according to the global usability score), on the vertical axis is the cumulative probability for each device to be the best option, among the best two options, among the best three options, and so on.

Figure 5. The thresholds for ranking DPIs usability by means of the estimated S-GUS scores: white points represent the median S-GUS while the bars indicate the 95% credibility interval; we calculated such statistics considering the GUS score of each device when the device was ranked 1st to 2nd, 3rd to 4th and 5th to 6th, respectively.
Discussion

The choice of the most convenient DPIs to prescribe is too frequently empiric in real life, being the role of effective criteria of choice often underestimated or neglected [5,13,15,16]. Usually, the patients’ skill of the operational manoeuvres leading to the proper inhalation of the drug from DPIs is usually inadequate and only occasionally checked before their prescription: it is regarded as a time-consuming procedure that also requires a prescribers’ specific knowledge [5,13]. This represents a crucial point particularly in long-term inhalation treatments, as in the case of persistent airway diseases (i.e. bronchial asthma and chronic obstructive airway disease – COPD), when the personalization of therapy is absolutely relevant in terms of patients’ adherence to treatments and then of outcomes to achieve. Even if DPIs presently available are highly effective, nonetheless it is known that patients are not able to use all of them equally well [5,6,21,24]. Moreover, the sophisticated engineering of present DPIs would increasingly require the contribution of advanced health professional roles or specific knowledge for teaching, training and supporting patients properly [6,7,32]. Despite these critical aspects, inhalation absolutely remains the route to prefer for delivering and assuming respiratory drugs effectively. Consequently, usability should be regarded as the most important parameter contributing to the motivated and convenient choice of the DPI to prescribe [19-21].

The term “usability” was sometimes merely used in the literature as a synonym of “ease of use”, or “intuitivity”, or “preference at glance”, even if usability itself does represent a much more complex and comprehensive concept than expected. Actually, usability mirrors the role of a much wider panel of variables underlying and affecting proper inhalation. In other words, usability reflects the influence of those factors depending of the patients’ point of view (i.e., preference, attractivity, etc.) [12,15], but also of those related to the effects of intrinsic technological characteristics of the DPIs, of which patients are usually almost completely unaware (i.e., emission velocity, lung deposition, dose consistency, etc.). The careful assessment of DPIs usability through quantitative and specific indices should then be increasingly favoured as a priority issue also in the aim of comparing and ranking their usability objectively.

Nevertheless, the great majority of recent investigational instruments has been mostly (or uniquely) developed in order to mainly focus patients’ intuitiveness, preference, and satisfaction related to DPIs [11,14,16-18]. Unfortunately, even if these instruments proved to be reliable tools for checking patients’ personal beliefs, the role of an independent assessment by a third-party expert observer (i.e., by an expert supervising nurse or doctor) was sometimes lacking, or not valued systematically [16,17]. In our center, two nurses, equally expert, motivated in educational programs, and familiar with the technical and the psychological aspects of the GUS Questionnaire, played this role. They were specifically dedicated to patients’ interviews, and to supervise, check, assess, and validate all patients’ procedures for inhalation. However, this figure could be generalizable and extended to any healthcare professional or other caregiver, but only after a specific training.

Furthermore, specificity and suitability represent the most important qualities, which should characterize the investigational instruments to use in clinical studies for investigating DPIs performances. If O-GUS allowed for the first time to discriminate DPIs in terms of their effective usability [24], the same extent of inter-DPIs differences was also equally documented with the S-GUS: not only in terms of patients’ perceptions and beliefs, but mostly in terms of number of attempts and of time required for achieving the first proper actuation of DPIs. In particular, data from the present study strongly confirm the good suitability of the novel S-GUS in clinical practice as they are super-imposable to those obtained with the extended O-GUS and prove the almost absolute correspondence between patients’ and nurse’s opinions.

The role of time to spend for checking one more DPIs objectively and to overcome the difficulties encountered by patients prior to their prescription still represents a key parameter in clinical practice. To emphasize that the novel S-GUS consents the ranking of DPIs in a very short time, thus highly favouring the regular check of patients’ DPIs usability in current clinical practice. Furthermore, the short-form questionnaire can be easily converted into a mobile application that represents one of the more useful instruments in medical practice.

Conclusions

DPIs were further confirmed to be characterized by different degrees of usability. S-GUS proved as a quite specific tool for assessing DPIs usability in short time. The availability of this suitable, reproducible, smart and quick operational tool should be regarded as a valuable progress in this particular field of investigation.

Acknowledgements

None.

References


Received for publication: 24 March 2020. Accepted for publication: 27 April 2020.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

©Copyright: the Author(s), 2020

Licensee PAGEPress, Italy

Multidisciplinary Respiratory Medicine 2020; 15:659