Handling missing items in the Exacerbations of Chronic Pulmonary Disease Tool

To the Editor:

The Exacerbations of Chronic Pulmonary Disease Tool (EXACT) is a 14-item, self-administered daily symptom diary designed to identify and characterise exacerbations of chronic obstructive pulmonary disease (COPD). It provides a reliable, valid and standardised measure of exacerbation symptoms, and is sensitive to changes during recovery [1]. Scores are expressed on a 100-point scale, with higher values indicating worse symptoms or health state. In addition, the EXACT-derived E-RS (EXACT Respiratory Symptoms) provides valid daily COPD symptom scores [2, 3]. Electronic administration is recommended and has several advantages, notably in preventing item omission [4]. However, the expense of electronic solutions may prove prohibitive, particularly in noncommercial studies, when a pen-and-paper version may be used instead. In this context, it is important to have a method to deal with missing items. This is yet to be established.

In psychometrically validated instruments with high internal consistency, such as the EXACT (Cronbach-α $\geqslant$0.9 [1, 5]), missing items may be imputed from the remaining answered items [6]. Used appropriately, this is preferable to list-wise deletion of incomplete records, which reduces power and risks introducing bias if data are missed not at random, and to substituting values from neighbouring records ("last observation carried forward" or "next value carried backward"), as this assumes symptoms are in steady state, which is unlikely during an exacerbation [6]. However, imputing items increases random error and, if items are missed systematically, may introduce bias. These factors limit the number and combination of items that can be imputed without excessively compromising reliability and accuracy. We sought to define the parameters under which this may be done by simulating item imputation on complete EXACT records from a recent study.

The study was an investigator-led, multicentre, randomised, double-blind, placebo-controlled trial (www.clinicaltrials.gov identifier number NCT01247870 and www.isrctn.com identifier number ISRCTN66148745). Its methods and results are detailed elsewhere [7]. Briefly, the trial tested metformin in 52 patients admitted to hospital for COPD exacerbations, primarily to establish its antihyperglycaemic effect. Secondary end-points included symptomatic recovery, as determined by the EXACT. Eligible patients were aged $\geqslant$35 years, had established COPD and had been admitted for an exacerbation with an expected inpatient stay $\geqslant$48 h. Participants completed the EXACT on paper every evening for 1 month, including in hospital. Guidance was provided by investigators in person during the inpatient phase and telephone support was available following discharge.

The effect of imputing items was simulated on EXACT diary records from the first 17 participants, representing all participants enrolled by the time of this analysis. This dataset comprised 361 EXACT diary records, of which 302 (84%) were complete. In the first simulation, one randomly selected item was deleted from each complete record and an imputed score substituted in its place. Imputed scores were calculated as the mean raw score from the remaining items, rounded to the nearest integer and capped at the maximum available for the item being imputed. The total imputed and actual raw scores were transformed to a 100-point linear scale for analysis and interpretation [4]. The degree to which systematic error (bias) was introduced by imputation was quantified by the mean difference between imputed and actual scores (imputed−actual). Random error was quantified by the standard deviation of the difference and 95% limits of agreement were calculated [8]. To stabilise the estimates, they were averaged from 500 iterations of the item-imputation simulations. The same procedure was adopted to evaluate the effects of imputing 2–6 randomly selected items.

To identify items susceptible to systematic omission due to changing symptoms or setting of care, omission rates were compared between the inpatient and outpatient phases of the study. Those items with differential omission rates were subjected to further analysis using complete EXACT records from all 52 trial participants, systematically imputing these items on 12 representative days (days 2–10, 15, 20 and 28, where day 1 was defined as the day of admission). Mean difference, standard deviation and 95% limits of agreement were calculated for all possible item-omission combinations for these three items.

Overall, the mean±SD difference between imputed and actual scores ranged from −0.06±1.12 to +0.26±4.42 points, depending on the number of items imputed. These differences and their 95% limits of agreement are illustrated in figure 1. Items that were omitted significantly more frequently in the inpatient phase were item 9 (breathlessness with personal care; omitted on 5% of inpatient days *versus* 1% of outpatient days), item 10 (breathlessness with indoor activities; 28% *versus* 10%, respectively) and item 11 (breathlessness without outdoor activities; 31% *versus* 15%, respectively) (p<0.001 in each case). For these items, the overall mean±SD difference (95% limits of agreement) was −0.6±0.8 (−2.2–+1.1) points when any one item was imputed, −1.2±1.6 (−4.2–+1.9) when any two items were imputed and −2.0±0.3 (−6.6–+2.8) when all three were imputed. This appeared stable over the 12 days on which this was analysed.

Our results suggest that, in general, imputation introduces negligible systematic bias in the total EXACT score, but random error increases progressively with the number of items imputed. The average intrapatient day-to-day variability in EXACT scores is ∼5 points [1] and we considered this a reasonable benchmark against which to set the tolerance margin for imputation. In general, the 95% limits of agreement are within this margin provided that no more than three items are imputed. Items 9–11 represent a special case, because their omission had a systematic component. Imputing these items generated bias towards underestimating the actual score. To keep the 95% limits of agreement within the 5-point tolerance margin, no more than two of these items should be imputed. In doing this, a mean bias of −1.2 points is generated. This is relatively small in comparison to the effect of an exacerbation (generally ⩾10 points [5, 9]) but it should be borne in mind during data interpretation. The high omission rates for items 9–11 among hospital inpatients may raise questions over their validity in this setting. However, while the method we have proposed can compensate for partial missing data, prospectively removing these items for inpatients would require a new validation process.

A limitation of these simulations is that, by necessity, they were performed on complete diary records. One cannot know whether imputation has the same effects where the participants have themselves elected to omit items. The pattern of omission suggests that this might occur where participants consider the questions inapplicable to their present condition. For example, the patients may judge that quantifying breathlessness outside the home may be inapplicable whilst they are in hospital. Compelled to answer, these patients may have responded differently to those who answered the question from the outset. That said, we stressed to participants in the trial the importance of answering all items on every occasion. Consequently, the dataset probably includes reasonable representation from participants who recognised the same incongruence but answered the question anyway.

In conclusion, we recommend that missing EXACT items may be imputed from the remaining answered items, provided no more than three items are imputed in total, including no more than two of items 9–11. The imputed score is calculated from the mean of the remaining item scores, rounded to the nearest integer and capped at the maximum score available for that item.
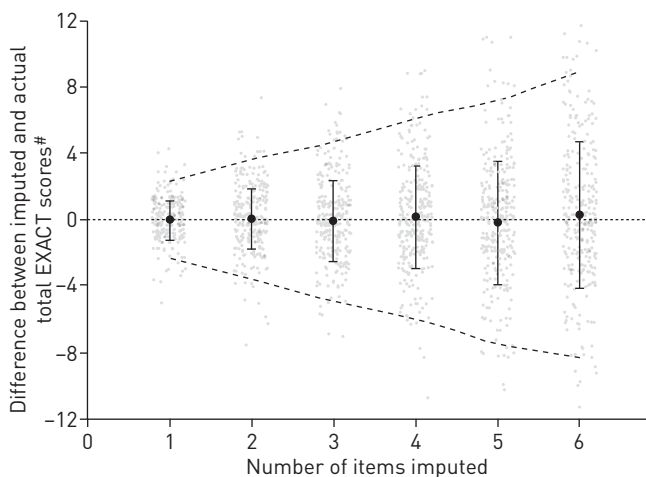

FIGURE 1 Difference between imputed and actual Exacerbations of Chronic Pulmonary Disease Tool (EXACT) scores, according to the number of items imputed. EXACT scores are expressed on a 100-point scale, with higher scores indicating worse symptoms or health state. Black circles denote the mean difference between actual and imputed scores, error bars denote the standard deviation of the difference and dashed lines indicate 95% limits of agreement. These estimates were generated from 500 iterations of random item-imputation simulations. Illustrative data points from one such iteration are denoted by grey spots. [#]: imputed−actual.

Andrew W. Hitchings, Emma H. Baker and Paul W. Jones
Institute for Infection and Immunity, St George's, University of London, London UK.

Correspondence: Paul W. Jones, Institute for Infection and Immunity, St George's, University of London, Cranmer Terrace, London, SW17 0RE, UK. E-mail pjones@sgul.ac.uk

References

1   Leidy NK, Wilcox TK, Jones PW, *et al.* Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *Am J Respir Crit Care Med* 2011; 183: 323–329.
2   Leidy NK, Murray LT, Monz BU, *et al.* Measuring respiratory symptoms of COPD: performance of the EXACT-Respiratory Symptoms Tool (E-RS) in three clinical trials. *Respir Res* 2014; 15: 124.
3   Leidy NK, Sexton CC, Jones PW, *et al.* Measuring respiratory symptoms in clinical trials of COPD: reliability and validity of a daily diary. *Thorax* 2014; 69: 443–449.
4   EXACT-PRO Initiative. The Exacerbations of Chronic Pulmonary Disease Tool (EXACT) Patient-Reported Outcome (PRO) user manual (version 6.0). Bethesda, Evidera, 2013.
5   Leidy NK, Murray LT, Jones P, *et al.* Performance of the EXAcerbations of Chronic pulmonary disease Tool patient-reported outcome measure in three clinical trials of chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2014; 11: 316–325.
6   Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res* 2014; 23: 440–459.
7   Hitchings AW, Lai D, Jones PW, *et al.* Metformin in severe exacerbations of chronic obstructive pulmonary disease: a randomised controlled trial. *Thorax* 2016 [in press DOI: 10.1136/thoraxjnl-2015-208035].
8   Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
9   Mackay AJ, Donaldson GC, Patel AR, *et al.* Detection and severity grading of COPD exacerbations using the exacerbations of Chronic Obstructive Pulmonary Disease Tool (EXACT). *Eur Respir J* 2014; 43: 735–744.