

ARTICLE OPEN

Device errors in asthma and COPD: systematic literature review and meta-analysis

Henry Chrystyn¹, Job van der Palen^{2,3}, Raj Sharma⁴, Neil Barnes^{4,5}, Bruno Delafont⁶, Anadi Mahajan⁷ and Mike Thomas^{8,9,10}

Inhaler device errors are common and may impact the effectiveness of the delivered drug. There is a paucity of up-to-date systematic reviews (SRs) or meta-analyses (MAs) of device errors in asthma and chronic obstructive pulmonary disease (COPD) patients. This SR and MA provides an estimate of overall error rates (both critical and non-critical) by device type and evaluates factors associated with inhaler misuse. The following databases from inception to July 23, 2014 (Embase[®], MEDLINE[®], MEDLINE[®] In-Process and CENTRAL) were searched, using predefined search terms. Studies in adult males and females with asthma or COPD, reporting at least one overall or critical error, using metered dose inhalers and dry powder inhalers were included. Random-effect MAs were performed to estimate device error rates and to compare pairs of devices. Overall and critical error rates were high across all devices, ranging from 50–100% and 14–92%, respectively. However, between-study heterogeneity was also generally >90% (I-squared statistic), indicating large variability between studies. A trend towards higher error rates with assessments comprising a larger number of steps was observed; however no consistent pattern was identified. This SR and MA highlights the relatively limited body of evidence assessing device errors and the lack of standardised checklists. There is currently insufficient evidence to determine differences in error rates between different inhaler devices and their impact on clinical outcomes. A key step in improving our knowledge on this topic would be the development of standardised checklists for each device.

npj Primary Care Respiratory Medicine (2017)27:22; doi:10.1038/s41533-017-0016-z

INTRODUCTION

Inhaled medications are fundamental to the treatment of asthma and chronic obstructive pulmonary disease (COPD), with inhaler devices being the principal route for administering such treatments.^{1, 2} Many different types of inhaler are available, but pressurized metered dose inhalers (pMDIs) and dry powder inhalers (DPIs) are the devices most commonly used for drug delivery in the treatment of asthma and COPD.^{1, 2} A large number of asthma and COPD patients do not use their inhaler devices correctly. Errors in device use may impact the effectiveness of the delivered drug and thereby lead to the sub-optimal control of asthma and COPD.^{3–6} It is therefore important to understand and quantify device-use errors so that patient interventions can be effectively introduced and new devices designed to avoid common errors.

The literature highlights the fact that the definitions of critical and non-critical errors, as well as the number and type of checklist steps, vary widely between different devices and studies. A critical error is one that may impact the effectiveness of the delivered drug and thereby lead to the sub-optimal disease control of asthma and COPD,^{3, 7} whereas a non-critical error is one of the checklist steps for a particular device that is not classified as critical.

A previous systematic review (SR) focused on errors with DPI devices only⁷ and another has found that there has been no change in the type and number of errors reported over the past 40 years.⁸ The present SR and MA was conducted to provide an estimate of error rates (i.e., the proportion of patients with at least

one error, critical and/or non-critical) by device type and to evaluate the factors associated with inhaler misuse, for example device and patient characteristics. In addition, the use of educational interventions designed to improve inhaler technique was investigated.

RESULTS

Search results

The search results for the SR are shown in the PRISMA flow diagram (Fig. 1). Overall, a total of 2519 citations were identified via database searching and a further 18 were identified through back-referencing of reviews and other relevant primary studies. After screening, 72 primary studies were extracted, all of which were included in the SR and 40 of these were selected for inclusion in the MA, based on the predefined criteria previously described (Fig. 1). Reasons for study exclusion are provided as Supplemental data (Appendix 1).

Study characteristics

The majority (54%) of the 72 identified studies included in the SR comprised patients with both asthma and COPD, while 32% and 14% were conducted in either asthma or COPD patients, respectively. Most (90%) of the mixed patient studies did not report data for the asthma or COPD subgroups separately.

¹Inhalation Consultancy Ltd, Yeadon, Leeds and Faculty of Human and Health Sciences, Plymouth University, Plymouth, UK; ²Department of Research Methodology, Measurement, and Data Analysis, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, The Netherlands; ³Department of Epidemiology, Medisch Spectrum Twente, Enschede, The Netherlands; ⁴GSK, Brentford, UK; ⁵William Harvey Institute, St Bartholomew Hospital and The London School of Medicine and Dentistry, London, UK; ⁶Exploristics Ltd, Belfast, UK; ⁷Bridge Medical, Richmond, UK; ⁸Primary Care and Population Sciences, University of Southampton, Southampton, UK; ⁹NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK and ¹⁰NIHR Wessex Collaboration for Leadership in Applied Health Research and Care (CLAHRC), Southampton, UK

Correspondence: Henry Chrystyn (h.chrystyn@gmail.com)

Received: 31 March 2016 Revised: 16 December 2016 Accepted: 1 February 2017

Published online: 03 April 2017

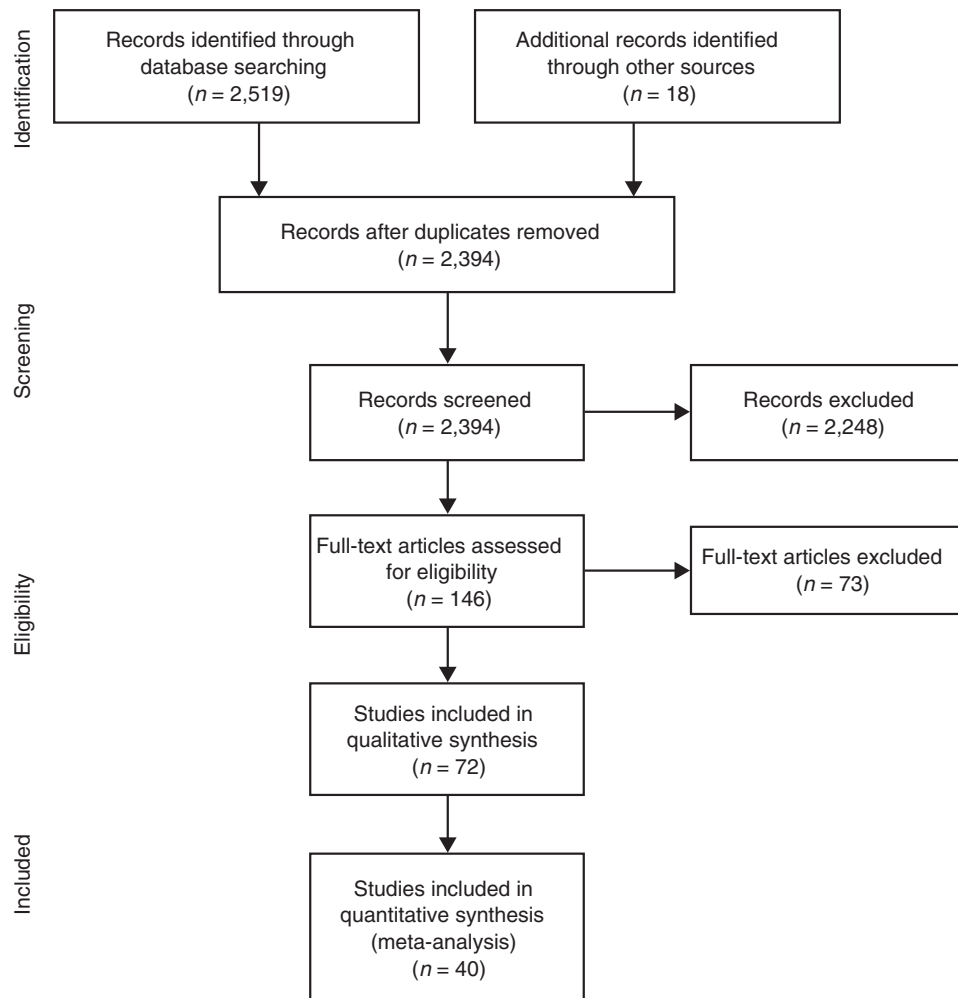


Fig. 1 PRISMA flow diagram describing the results of the systematic review

Fifty-four observational studies were retrieved. Of these, 40 were of a cross-sectional design and the remaining studies were prospective. Only 18 randomised controlled trials (RCTs) were identified, of which 11 were crossover studies and the other seven were of a parallel group design. Approximately 80% of the observational studies included in the MAs were conducted in an outpatient population and with patients primarily utilising their existing inhaler device. Baseline data were obtained prior to device training for most of the RCTs.

The majority of extracted studies evaluated MDI ($n = 29$) or DPI ($n = 32$) inhalers and reported either the proportion of patients with any error (critical plus non-critical) or those with any critical error. Amongst the DPI studies, most involved Turbuhaler® ($n = 17$) or Diskus® ($n = 15$) devices. Fifty-four studies were conducted in patients who were regular inhaler users and assessed the inhalation technique that patients had been employing on a regular basis.

Due to limited data availability, MAs of any device comparisons were not feasible for the: i) overall error frequency (cross-sectional studies and RCTs) and ii) critical error rates for the RCTs.

Most of the studies included in the overall analysis ($n = 72$) were conducted in the USA ($n = 13$), the Netherlands ($n = 10$) and the UK ($n = 10$).

Amongst all of the studies, about two-thirds reported that the assessor was trained in the inhalation technique of the device under assessment. The assessors included pharmacists, physicians/general practitioners (GPs)/specialists, students,

investigators, research assistants and technicians. In the majority (>90%) of cases, inhalation technique was assessed utilising author-validated, existing checklists. Furthermore, a variety of checklists was used for the same device across different studies. These checklists were highly variable and differed in both the number of steps and their definition(s) of these steps. Errors arose from failure to correctly complete the relevant checklist steps for a particular device.

Supplementary Table 2 details the characteristics of all studies that reported overall and critical error rates. The number of studies included in the MA was lower (<72) because studies that reported, i) an error frequency using a definition other than that previously defined, ii) pooled data, or iii) error rates for each individual step, but not cumulatively, were excluded.

Overall error rate (critical and non-critical)

Across the devices, error rates appeared common with approximately 50–100% of patients experiencing at least one error. The pooled summary results for the MDI devices estimated an overall error frequency (RE model for all studies) of 86.8% [95% CI 79.4–91.9] of patients with at least one error (Fig. 2a). However, there was a high level of heterogeneity between the studies (98.5% I-squared statistic). Compared with MDIs, relatively few studies assessed the overall error rates for DPI devices. The pooled summary results for DPIs estimated an overall error frequency (RE model for all studies) of 60.9% [95% CI 39.4–79.0] of patients with

at least one error, with a high level of between-study heterogeneity (99.0% I-squared statistic; Fig. 3a).

The frequency of overall error rates for individual devices is shown in Table 1. None of the studies assessed the overall or critical error rates for the Breezhaler®, Easyhaler®, Ellipta®, Elpenhaler®, Genuair®, Nexthaler®, or Novolizer® devices; devices with zero studies have been excluded from the table. An MA of overall error frequency for Turbuhaler® and Diskus® in prospective and cross-sectional studies is shown as Supplemental data (Appendix 2, Supplementary Figure 1). The most common overall errors by device are detailed in Table 2. A sensitivity analysis of the overall error frequency for MDIs, conducted to assess any bias due to the inclusion of industry-sponsored studies, gave a similar finding (87.6% [95% CI 79.4–92.9]; RE model for all studies) to the pooled summary (Appendix 3; Supplementary Figure 3).

Critical error rates

Across the devices, critical error rates appeared common with approximately 14–92% of patients experiencing at least one critical error. The frequency of critical error rates for individual devices is shown in Table 3. The pooled summary results for MDIs estimated a critical error frequency of 45.6% [95% CI 26.0–66.6] of patients with at least one critical error ($n = 10$ studies), however, the data were highly heterogeneous (98.4% I-squared statistic) (Fig. 2b). The critical error rates for the DPIs were highly variable for each device: Aerolizer® ($n = 4$ studies) 14.2% [95% CI 11.0–18.1], Diskus® ($n = 9$ studies) 20.8% [95% CI 13.7–30.2], Turbuhaler® ($n = 10$ studies) 40.1% [95% CI 28.6–52.9], and Handihaler® ($n = 3$ studies) 42.4% [95% CI 28.8–57.1]. The pooled summary results for DPIs estimated a frequency of 28.4% [95% CI 22.0–35.8] of patients with at least one critical error (96.2% I-squared statistic; Fig. 3b). The between-study variability was high (>90% I-squared statistic) for the Diskus® and Turbuhaler® devices. The heterogeneity was lower for the Aerolizer® (44.3% I-squared statistic) and Handihaler® (58.4% I-squared statistic) devices but there were fewer studies available for inclusion, so the between-study variability may be under-estimated.

An MA of critical error frequency for Turbuhaler® and Diskus® in prospective and cross-sectional studies is shown as Supplemental data (Appendix 2, Supplementary Figure 2). MA using the RCTs was only feasible for the Diskus® inhaler. The pooled critical error estimate was 18.0% [95% CI 8.2–35.1] and the between-study heterogeneity was relatively high (84.2% I-squared statistic).

There were insufficient studies available (≤ 1) to enable quantitative analysis for the Autohaler®, Breezhaler®, Easyhaler®, Ellipta®, Elpenhaler®, Genuair®, Nexthaler® and Novolizer® devices.

Differences in error rates between devices

There was only sufficient evidence to conduct MAs of the critical error frequency for the observational studies.

Critical errors

The only significant results were for the comparison between the Diskus® and Turbuhaler® devices, odds ratio (OR): 2.90 [95% CI 1.41–5.96] ($n = 9$ studies) indicating fewer errors when using the Diskus® device, although the heterogeneity between studies was still very high (93.9% I-squared statistic). The pooled OR for critical error frequency favoured Turbuhaler® over MDI (OR: 1.76 [95% CI 0.53–5.85], $n = 7$ studies; 93.9% I-squared statistic) and Diskus® vs. MDI (OR: 5.01 [95% CI 0.87–28.69], $n = 6$ studies). However, in both cases, the heterogeneity between studies was >98% (I-squared statistic).

Trend differences were observed for the following device comparisons for critical errors: Turbuhaler® vs. MDI (OR: 1.76 [95% CI 0.53–5.85]), Diskus® vs. MDI (OR: 5.01 [95% CI 0.87–28.69]), and

Turbuhaler® vs. Handihaler® (OR: 1.09 [95% CI 0.47–2.53]), but the results were non-significant.^{3, 11, 16, 18, 25, 28, 30, 31, 33–39}

A sensitivity analysis, conducted to assess any bias due to the inclusion of industry-sponsored studies, provided similar results (Appendix 3; Supplementary Figures 4–6).

Impact of patient and study characteristics

The meta-regression analysis showed no significant findings for different baseline characteristics. However, a qualitative review of the extracted studies that analysed data according to, i) patient, ii) disease, and iii) other characteristics, and assessed the association of these factors with the likelihood of making a device error was conducted. A total of 37 primary studies assessed predictors of inhalation technique errors.

Common patient characteristics that were reported to impact device error rates included: i) age ($n = 15$ studies, with patients in older age groups reporting more errors compared with younger patients);^{3, 11, 16, 18, 25, 28, 30, 31, 33–39} ii) education ($n = 9$ studies)^{11, 16, 22, 24, 28, 31, 35, 36, 40} with a higher level of education being associated with fewer errors; iii) COPD diagnosis ($n = 6$ studies);^{15, 23, 27, 28, 31, 41} iv) gender (females reporting a higher error frequency);^{28, 35, 42} v) socioeconomic status (low income was associated with a higher frequency of errors); and vi) relationship status, number of comorbidities and disease severity ($n = 1$ study each).^{34, 43}

In studies that recruited patients with asthma or COPD, there was a higher error frequency in patients with COPD ($n = 6$ studies).^{15, 23, 27, 28, 31, 41} In the four studies that included only COPD patients,^{15, 22, 26, 29} two reported the overall error frequency (>93% of patients had at least one error,^{15, 22} MDI and Aerolizer® devices). In the two studies reporting critical error frequency,^{26, 29} 93% and 76% of patients were reported to have at least one critical error, respectively. However, the sample size was relatively small ($n < 30$) in three of the studies. Of the eight studies that recruited only asthma patients,^{3, 4, 10, 11, 13, 17, 23, 24} four reported an overall error frequency of <75%. Critical error frequency was not reported.

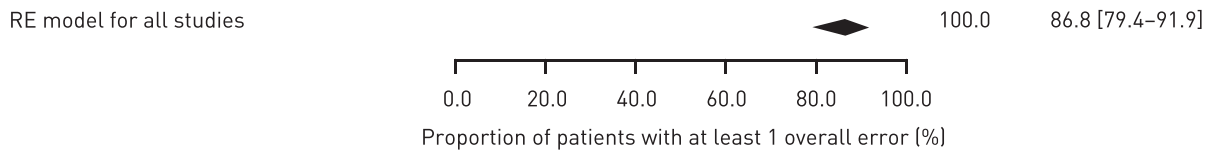
Inhaler-related characteristics reported to impact device error rates included prior training on device use,^{3, 11, 26, 28, 30, 31, 34, 36, 40, 41} and duration of device use and use of multiple inhalers ($n = 3$ studies each).^{25, 29, 41} Receipt of prior training on device use was predictive of a lower error frequency compared with no prior training; further, a longer duration of device training compared with shorter duration of training, and receipt of a practical demonstration of correct inhalation technique were factors also associated with lower error frequency rates. A longer duration of device use was predictive of a higher error rate, compared with those who had received their inhaler devices more recently. Use of multiple inhalers also predicted a higher error frequency.

Other characteristics that were reported to affect device error frequency included regular clinic visits,^{4, 28} polypharmacy,^{18, 41} and uncontrolled disease.^{4, 31} ($n = 2$ studies each). No trend was observed with regards to patient setting in either overall or critical error frequency.

There was a large variability in the sample size of the included studies. In general, studies with a smaller sample size reported a higher frequency of device errors, particularly critical errors.^{26, 29, 30} There was also a large variation in the number of steps assessed for the different inhaler devices, with studies including between three steps⁹ and 12 steps⁵ for overall error assessment. The relationship between number of steps and error rate was investigated by ordering forest plots of error rates by the number of steps in each study. There was a trend towards higher error rates with a larger number of steps, however no consistent pattern could be observed.

a

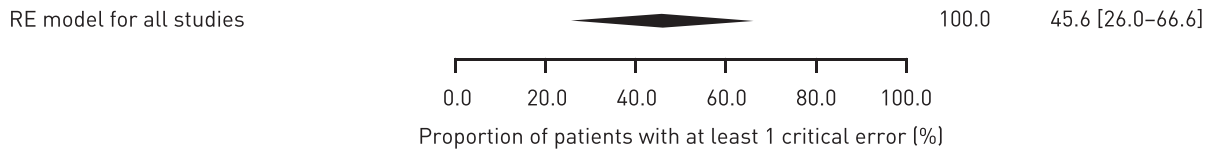
Study	N steps	N patients	Weight (%)	Prop. (%) [95% CI]
Sarvis 2004 ⁹	3	33	4.2	93.9 [78.8–98.5]
Hilton 1990 ¹⁰	4	262	6.0	55.0 [48.9–60.9]
Liard 1995 ¹¹	5	668	6.0	77.8 [74.5–80.8]
Allen 1986 ¹²	6	30	4.7	90.0 [73.2–96.7]
Al-Jahdali 2013 ¹	6	414	6.0	45.7 [40.9–50.5]
Shrestha 1996 ¹³	7	125	5.8	79.2 [71.2–85.4]
Molimard 2003 ¹⁴	7	552	6.0	76.1 [72.3–79.5]
Souza 2009 ¹⁵	7	67	4.7	95.5 [87.0–98.5]
Vargas 2013 ¹⁶	7	191	3.3	99.5 [96.4–99.9]
Zainudin 1990 ¹⁷	8	93	5.8	62.4 [52.1–71.6]
Horsley 1988 ¹⁸	9	86	5.2	94.2 [86.8–97.6]
Larsen 1994 ¹⁹	9	501	5.9	89.2 [86.2–91.7]
Erickson 1998 ²⁰	9	159	5.5	94.3 [89.5–97.0]
Plaza 1998 ²¹	9	746	6.0	91.0 [88.7–92.9]
Al-Hassan 2009 ²²	9	100	2.3	99.5 [92.6–100.0]
Epstein 1979 ²³	11	130	5.7	89.2 [82.6–93.5]
Giraud 2002 ⁵	12	3,955	6.0	70.6 [69.1–72.0]
Hashmi 2012 ²⁴		215	5.9	83.7 [78.2–88.1]
Arora 2014 ²⁵		70	5.0	94.3 [85.7–97.8]



I² (total heterogeneity / total variability): 98.46%
 Test for heterogeneity:
 Q(df = 18) = 510.5664, p<0.0001

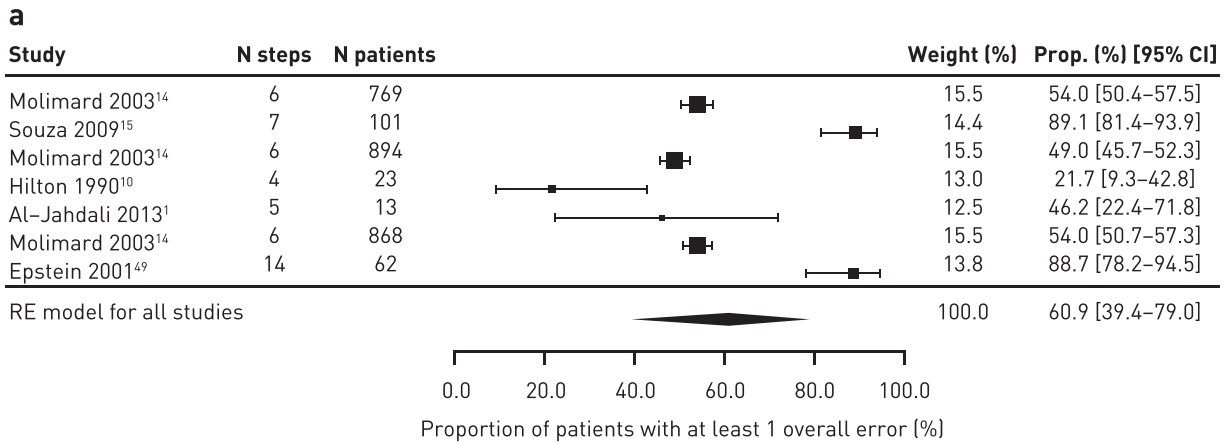
b

Study	N steps	N patients	Weight (%)	Prop. (%) [95% CI]
van der Palen 1995 ²⁶	3	25	9.7	76.0 [55.8–88.8]
Molimard 2003 ¹⁴	3	552	10.9	28.1 [24.5–32.0]
Khassawneh 2008 ²⁷	3	193	10.8	74.6 [68.0–80.3]
Allen 1986 ¹²	4	30	10.1	40.0 [24.3–58.1]
Hesselink 2001 ²	4	40	10.2	30.0 [17.9–45.7]
Melani 2004 ²⁸	4	866	10.9	23.9 [21.2–26.9]
Batterink 2012 ²⁹	4	14	6.8	92.9 [63.0–99.0]
Rootmensen 2010 ³⁰	5	32	9.8	81.2 [64.1–91.3]
Melani 2011 ³¹	5	843	10.8	12.0 [10.0–14.4]
Ho 2004 ³²	8	39	10.0	17.9 [8.8–33.1]

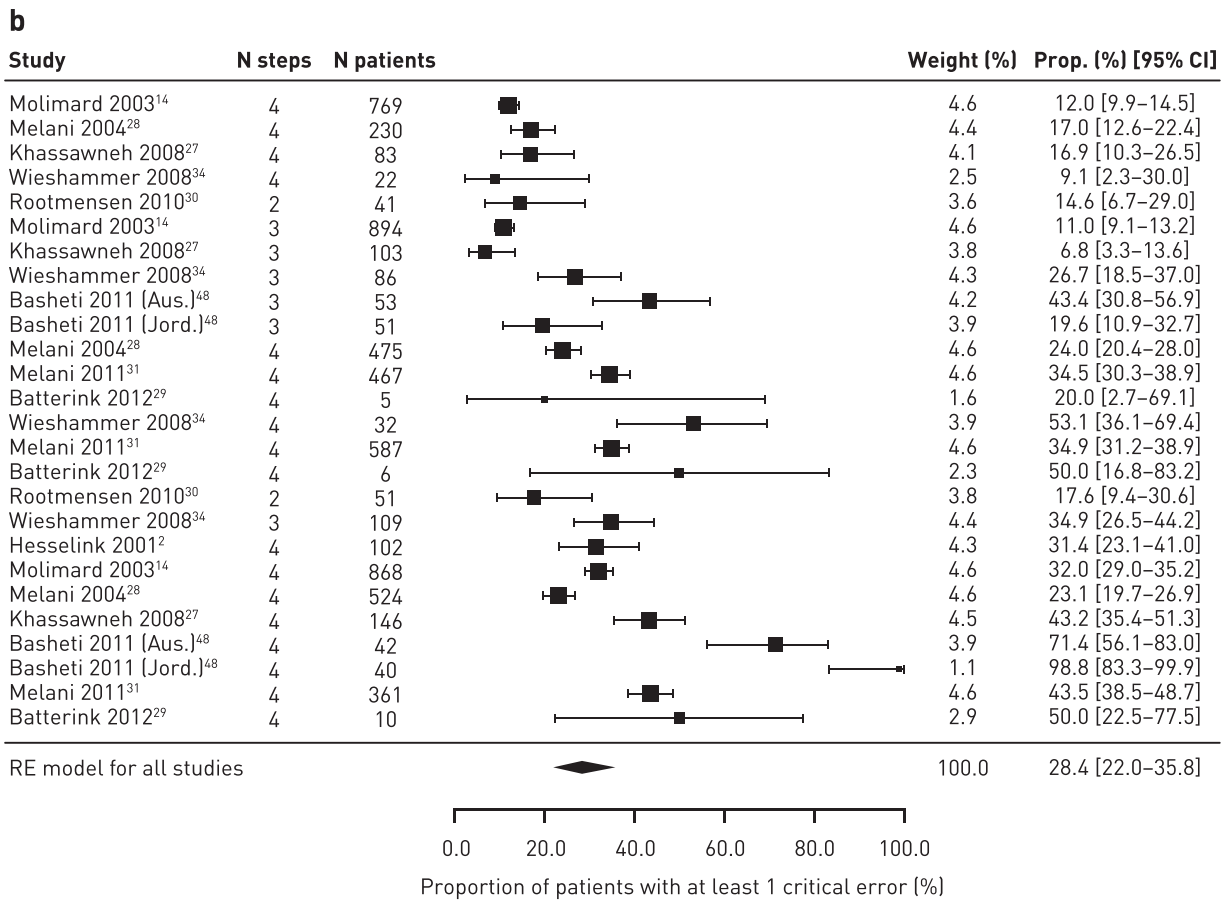


I² (total heterogeneity / total variability): 98.44%
 Test for heterogeneity:
 Q(df = 9) = 316.9279, p<0.0001

Fig. 2 Meta-analysis of the overall error rate frequency (a) and critical error rate frequency (b) for MDIs in prospective/cross-sectional studies



I² (total heterogeneity / total variability): 98.98%
 Test for heterogeneity:
 Q(df = 6) = 76.1824, p<0.0001



I² (total heterogeneity / total variability): 96.21%
 Test for heterogeneity:
 Q(df = 25) = 415.0933, p<0.0001

Fig. 3 Meta-analysis of the overall error rate frequency (a) and critical error frequency (b) for DPIs in prospective/cross-sectional studies

Table 1. Frequency of overall error with devices (quantitative and qualitative estimates)

Device	Cross-sectional studies			RCTs		
	Number of studies	Pooled estimate [95% CI] (I-squared)*	Range of error frequency	Number of studies	Pooled estimate [95% CI] (I-squared)*	Range of error frequency
MIDI	19	86.8% [79.4–91.95] (98.5%)	45.7–100.0%	2	89.5% [75.7–95.9] (60.0%)	85.1–94.2%
MIDI with spacer	3	52.0% [24.3–78.6] (87.3%)	31.6–78.0%	0	N/A [‡]	N/A [‡]
BA-MIDI (Autohaler [®])	2	57.7% [52.9–62.5] (84.9%)	54.9–59.9%	0	N/A [‡]	N/A [‡]
Turbuhaler [®]	4	55.5% [24.5–82.7] (93.6%)	22.0–88.7%	3	73.6% [55.6–86.1] (92.1%)	54.0–82.2%
Diskus [®]	1	N/A [‡]	49.0%	3	62.4% [51.9–72.0] (61.6%)	50.0–68.3%
Aerolizer [®]	2	75.1% [31.0–95.3] (97.2%)	54.0–89.1%	0	N/A [‡]	N/A [‡]
Handihaler [®]	0	N/A [‡]	N/A [‡]	1	N/A [‡]	66.7%

N/A Not applicable
* I-squared statistic indicates the percentage of variance that is attributable to between-study heterogeneity. The higher the percentage, the greater the heterogeneity. [‡]Not applicable as number of studies was ≤ 1

DISCUSSION

No SR or MA of this type has been previously published. The SR conducted by Lavorini and colleagues⁷ was limited as it focused on DPI devices only and an MA of device errors was not undertaken. The present SR of the existing data provides a valuable and timely assessment of the quality of the existing evidence base surrounding device errors. Despite limitations of the data, it can be seen that both the overall and critical error rates are reported to be high across all devices, ranging from 50–100% and 14–92%, respectively. Although there were very limited data on error rates and symptom control/long-term outcomes, one might hypothesise that correct use of the device is fundamental to the efficacy of the drug, and in this case, the reported error rates and critical error rates may result in sub-optimal treatment and disease control.

There were insufficient high-quality data to draw definitive conclusions about the comparative error frequency between different devices. Meta-regression analyses of patient and study characteristics were inconclusive. However, previous studies have reported associations between certain patient and study characteristics and device error rates (although these data were not quantified). Some studies have reported that older age^{31, 33, 35} and female gender^{35, 44} are associated with higher device error rates. Other socio-economic factors have previously been reported to influence device error frequency. A higher level of education has been reported to be associated with a lower frequency of errors.^{25, 31, 36} Additionally, a higher frequency of errors was found in patients with COPD than those with asthma.^{15, 28} Other factors reported in the literature to impact error frequency were receipt of prior training and type of training,³⁶ duration of device use²⁹ and use of multiple inhalers.^{27, 30} This may be because patients with a longer duration of device use are likely to have only received training/instruction when the device was first prescribed. The higher error rate reported with multiple inhaler usage may have been due to the higher burden and confusion associated with the use of different devices.

Limitations of the available data

Given the importance of this area of research, the number of publications focusing on device errors is relatively low, especially when compared with the overall volume of clinical research publications in COPD and asthma. Additionally, there was very little information on the association between device error rates and clinical outcomes. From the available publications, the overall quality was moderate for the majority of studies, with few high-quality studies^{14, 27, 30, 31, 33} (Supplementary Table 2). Whilst it is possible to draw qualitative learning from across the studies, the lack of consistency in studying device errors means that the MAs have to be interpreted with caution.

There were several potential sources of heterogeneity between the different studies including: i) differences in disease diagnosis; ii) heterogeneity due to varying study types; iii) large variability in the level of training received by patients; iv) variability in device step checklists (one of the most important limitations); v) variability in assessors' technique; vi) the subjectivity of each assessment; vii) heterogeneity due to patient-introduced bias (the Hawthorne effect); viii) studies not specifically designed to assess error rates between devices. This inter-study heterogeneity and lack of consensus around error rates and the types of errors associated with different devices may have an impact in the clinic, with healthcare professionals unclear about which inhaler to prescribe for their patients.

In terms of diagnosis, the majority of studies recruited both asthma and COPD patients and did not provide information on the sub-groups according to disease type.^{25, 27} Moreover, there was no validation of the disease diagnoses reported in the papers. Amongst the cross-sectional studies, patients were generally

Table 2. Most common overall errors reported for each device

Device	No exhalation before inhalation (% frequency—range)	Number of studies	Not holding breath for a few secs after inhalation (% frequency—range)	Number of studies	Not using a proper seal around mouthpiece (% frequency—range)	Number of studies
MDI	10.2–60.2	14	46.7–76.7	18		
MDI with spacer	12.1–73.9	2	22.8–79.7	3	0–28.0	5
BA-MDI (Autohaler®)	22.3–23.0	2	30.2–38.7	2		
Turbuhaler®	16.0–75.0	12	6.0–77.1	15	0–21.7	5
Diskus®	20.6–65.8	9		12	0–15.1	10
Aerolizer®	24.5–77.1	4	18.7–77.1	5		
Handihaler®	25.0–77.1	3	24.7–77.1	4		

observed using their regular inhaler.^{14, 31} However, in the RCTs, patients were assessed in a controlled manner and included normally recruited device-naïve participants only.^{45, 46}

In the cross-sectional studies, device technique was assessed in patients without any study protocol-specific training or instruction prior to study entry. Moreover, the majority of the studies included in the MA specified a “lack of training regarding device use” as the primary reason for device mishandling.^{14, 31} Patients who received training had usually received it at the time of first prescription of the device and did not receive any additional follow-up training or further assessment of their inhaler technique.³¹ Additionally, the medical personnel responsible for teaching the correct inhaler technique were reported to be lacking in basic training.⁶ Inhalation technique was assessed by a variety of assessors including trained pharmacists,^{22, 27} respiratory specialists,⁴⁰ GPs,^{3, 14} or others (including trained assistants to the physician, intern students, or laboratory staff). Additionally, not all of the studies provided information regarding the assessors.

There was also variability in the number of steps, the actual details of the device checklist steps and in the definitions of critical and non-critical errors. For example, for the Diskus® device, the number of overall and critical errors possible varied between 5–12 and 2–4, respectively^{29–31, 47, 48} and the frequency of errors for the Turbuhaler® was 4–14 and 3–5, respectively.^{2, 10, 30, 49} There was a trend towards higher error rates with a larger number of steps, however no consistent pattern was observed. The definition of critical errors used in the majority (approximately 90%) of studies (e.g., “the proportion of patients with an error for a step that is deemed necessary for the adequate delivery of the drug to the lungs”) is also highly ambiguous (i.e., adequate could mean anywhere between 20–100% drug delivery).

There may also be heterogeneity due to patient-introduced bias, i.e., the Hawthorne effect. This is when the patient is being observed for the technique as part of the study/studies where there is a high probability that the patient will try to use the inhaler in the best possible way.³⁷ This may not result in a true reflection of their daily use and leads to an under-estimation of the error frequency. A few studies specified that inhaler technique demonstration took place in an empty room and that the videotaped observations were assessed by nurses.⁵⁰

Finally, although the quality of the data included in these studies was good (Supplementary Table 2), the studies included in these comparisons were largely cross-sectional or prospective in nature and were not designed to compare error rates between different devices. Utilising them in this way necessitates caution. Additionally, a number of the studies had a relatively small sample size, for example, the Batterink study²⁹ showed significant differences in error rates for comparisons between MDIs and Turbuhaler® and Diskus® devices, but only included ten and five patients, respectively for the latter two devices.

Limitations of the review methodology

There are a number of limitations associated with this MA. Firstly, studies that reported error frequency using a definition other than that previously defined were not included in the analysis. This excluded approximately 10% of the studies. Secondly, studies that reported error rates for each individual step, but not cumulatively, were described qualitatively only (approximately 10% of studies). Additionally, studies that reported pooled data, i.e., reported for all DPI devices, were not included in the analysis ($n = 3$ studies).

Main implications for clinical practice and research

Clinical research. There is a need to standardise the definitions of non-critical and critical device errors and their assessment, as well as improve clarity on the clinical importance of each error. Indeed, the literature highlights that definitions of critical and non-critical errors can vary substantially between different studies. It is essential that the wider scientific community reaches a consensus regarding error terminology for the different inhaler devices and develops a standardised checklist for each device, similar to that which is available in the Netherlands.⁵¹ This will not only be useful for standardising the conduct of future clinical research but also will also provide a valuable clinical tool and enable comparison of devices across future studies.

It should be noted that a number of errors are common to both the MDI and DPI devices, e.g., “exhaling before inhalation” and “holding breath for a few seconds after inhalation”. “Ensuring a proper seal around mouthpiece” is common for MDI with spacer, Diskus® and Turbuhaler® devices (Table 2). It was not possible to identify common errors for Autohaler®, Breezhaler®, Elpenhaler®, Genuair®, Nexthaler® or Novolizer® due to limited data. These steps may provide the basis for the identification and refinement of errors. However, there appears to be more inconsistency in the definition of critical errors.

Once the step checklists and critical/non-critical errors have been standardised, there is a need to conduct more systematic clinical research in this area. There is currently insufficient evidence to be able to determine whether there are any differences in error (non-critical and critical) rates between different inhaler devices. Prospective clinical studies in inhaler-naïve patients, using more objective device training, are required in order to address this issue.

During routine clinical practice, inhaler technique errors may be compensated by increasing the medication dose if disease control is sub-optimal, but this has not yet been systematically studied. Conducting a study to assess low medication doses where the impact of critical device errors is likely to be more pronounced would provide a useful approach to assess the impact of various inhalation steps and errors on patient outcomes. Another approach would be to assess treatments that are available across

Table 3. Frequency of critical error with devices (quantitative and qualitative estimates)

Device	Cross-sectional studies			RCTs		
	Number of studies	Pooled estimate [95% CI] (I-squared)*	Range of error frequency	Number of studies	Pooled estimate [95% CI] (I-squared)*	Range of error frequency
MDI	10	45.6% [26.0–66.6] (98.4%)	12.0–93.0%	0	N/A [‡]	N/A [‡]
MDI with spacer	3	8.9% [0.9–50.5] (93.7%)	2.6–47.0%	0	N/A [‡]	N/A [‡]
BA-MDI (Autohaler [®])	1	N/A [‡]	11.0%	0	N/A [‡]	N/A [‡]
Turbuhaler [®]	10	40.1% [28.6–52.9] (95.7%)	18.0–100%	1	N/A [‡]	26.0%
Diskus [®]	9	20.8% [13.7–30.2] (93.6%)	6.8–43.0%	3	18.0% [8.2–35.1] (84.2%)	8.0–31.9%
Aerolizer [®]	4	14.2% [11.0–18.1] (44.3%)	9.1–16.8%	0	N/A [‡]	N/A [‡]
Handihaler [®]	3	42.4% [28.8–57.1] (58.4%)	35.0–53.1%	0	N/A [‡]	N/A [‡]

N/A Not applicable
* I-squared statistic indicates the percentage of variance that is attributable to between-study heterogeneity. The higher the percentage, the greater the heterogeneity. [‡]Not applicable as number of studies was ≤ 1

a number of different devices at microgram-equivalent doses, and to conduct real-world studies in COPD and asthma patients (individually and combined).

Finally, as the outcomes data are very limited, there is a need to investigate the links between different critical device errors and long-term, clinical effectiveness (patient outcomes), resource use and adherence rates.

Clinical practice. There has been awareness of the problem of device errors for over 40 years⁸ but this issue has still not been resolved. The high error rates observed across inhalation devices in this study suggests that more time should be invested by healthcare professionals in educating/training patients on how to operate their inhalers correctly. There are several factors that need to be addressed, including the requirement for standardised training of healthcare professionals and patients for the different inhaler devices and regular re-evaluation of inhaler technique and mastery. Existing guidelines^{3, 4, 49} provide targets for device training but these have still not been achieved. The development of standardised training protocols and schedules for the individual inhaler devices may aid this process. There is also a need to assess the ease of training and continued mastery across the different devices.

Once sound research techniques are available, this could lead to improvements in clinical practice whereby training is standardised and conducted on an ongoing basis, with regular re-assessment of patient device handling.

CONCLUSIONS

This SR and MA highlight the relatively limited body of evidence assessing device errors that is currently available, given the importance of this issue. From the available data in the literature, it is apparent that patients are not operating their inhalers correctly. Overall and critical error rates appear to be high across all of the devices assessed: approximately 50–100% and 14–92%, respectively. However, the high level of heterogeneity between studies prevents any definitive conclusions being drawn. There is currently insufficient evidence to be able to determine whether there are any differences in error (non-critical and critical) rates between different inhaler devices. Furthermore, there are limited available data assessing the impact of device errors on clinical outcomes. There is a need to develop and utilise consistent definitions of non-critical and critical device errors and to develop standardised checklists for each individual device in order to facilitate future clinical research and enable comparability between studies. The development of standardised training protocols for the individual inhaler devices will also aid this process.

METHODS

Search strategy

Excerpta Medica Database [Embase[®]], Medical Literature Analysis and Retrieval System Online [MEDLINE[®]], MEDLINE[®] In-Process (to ensure that non-indexed citations were retrieved) and the Cochrane central register of controlled trials (CENTRAL) databases were searched from inception to July 23, 2014. MEDLINE[®] and Embase[®] were searched using embase.com and CENTRAL and MEDLINE[®] In-Process were searched using Cochrane library and Pubmed.com interfaces, respectively. The search strategy is summarised in Supplementary Table 1.

Inclusion/exclusion criteria

Studies in adult males and females with asthma or COPD (all severity grades) were included in the analysis. The devices of interest were MDIs (both with and without spacers; chlorofluorocarbon/hydrofluoroalkane [CFC/HFA] propellant; breath-actuated [Autohaler[®] and EasiBreathe[®]] and Respimat[®] [soft mist]) and DPIs (Aerolizer[®], Breezhaler[®], Diskus[®]/Accuhaler[®], Easyhaler[®], Ellipta[®], Elpenhaler[®], Genuair[®], Handihaler[®], Nexthaler[®], Novolizer[®] and Turbuhaler[®]).

Data extraction

All citations (titles and abstracts) were screened for eligibility against the pre-specified inclusion/exclusion criteria. Full publications of the included citations were then reviewed for eligibility. All the citations excluded at the title/abstract or full-text screening stage were coded and the reasons for exclusion recorded.

Systematic review

Studies identified from the full-text screening stage that evaluated the number of patients with at least one overall error (critical plus non-critical) or at least one critical error, or assessed error frequency at each step for a specific device, were included in the SR.

Meta-analyses

Studies reporting at least one overall or critical error were included in the MA. For RCTs and prospective observational studies, the baseline error frequency prior to any study-related training or instruction was included in the analyses. Some of the device data were excluded from the MA because they did not meet criteria for the minimum number of studies and/or patients.

Quality assessment

The quality of all the included studies was assessed. The quality of RCTs was determined using the criteria published by the Cochrane Collaboration, with respect to different types of bias, and the quality of cross-sectional and observational studies was evaluated using the relevant Newcastle and Ottawa scale [2008]. Decisions about the estimated risk of bias were used to help evaluate heterogeneity between the studies.

Statistical analysis

Data from cross-sectional studies and RCTs were analysed separately.

For all quantitative analyses, data were evaluated for the proportion of patients with at least one overall error and those with at least one critical error. Studies that reported error frequency using a definition other than that previously defined and studies that reported pooled data, i.e., reported for all DPI devices, were not included in the analysis. Additionally, studies that reported error rates for each individual step, but not cumulatively, were described only qualitatively. Data were excluded from the analysis if there were fewer than five patients using a particular device within a study. MDI devices were categorised into two different sub-groups: i) MDI alone, ii) MDI with spacer.

A random-effects model was used for all the MAs in order to account for between-study heterogeneity.

Meta-analyses of device error frequency for each device/device type

All studies selected for analysis, and providing data for an individual device, were included in the MAs. The overall device error frequency was summarised for each device/device type using a restricted maximum likelihood (REML) random-effects MA. These analyses were performed when at least two studies provided adequate data for the same device. It should be noted that heterogeneity may be under-estimated when only a small number of studies were available.

Meta-analyses of device error frequency by sub-groups

The MAs described above were repeated by sub-group, i.e., i) diagnosis of asthma or COPD (asthma-only study, COPD-only study, mixed asthma/COPD study) and ii) previous device use (device-naïve patients, experienced device users, and a mix of naïve and experienced users).

Meta-analyses for comparison of pairs of devices/pairs of device types

Data from all the studies that provided error frequencies for the same two devices (within the same study) were included in the MAs. Pairs of devices were compared using a REML random-effects MA provided that a minimum of five studies for each comparison was available.

Meta-regression of device error frequency for each device/device type

Meta-regression is a tool for exploring the association of patient characteristics with outcomes of interest, thereby investigating sources of heterogeneity. Meta-regression aims to discern whether a relationship exists between an outcome measure and explanatory variables.

In order to assess the impact of baseline characteristics on the device error frequency, baseline characteristics reported in the studies were incorporated as regression factors into the REML random-effects MA model. Where available, the following baseline characteristics were considered for inclusion in the meta-regression analysis: i) population mean age; ii) proportion of current smokers; iii) proportion of males. However, as no significant findings were observed, these data are not reported.

Sensitivity analysis

A sensitivity analysis was also conducted to assess any potential bias resulting from the inclusion of industry-sponsored studies. In this analysis, data for inhaler devices that were products of the pharmaceutical company sponsoring the clinical trial(s), were excluded from the meta-analyses. The results of the sensitivity analysis were compared to the original analysis where all relevant studies were included.

ACKNOWLEDGEMENTS

Editorial support in the form of development of the draft outline and manuscript first draft in consultation with the authors, editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copy editing, fact checking, referencing, and graphic services was provided by Julie Wilson of Bridge Medical, Richmond, UK, and was funded by GSK. Additional editing and formatting prior to submission was carried out by Angela Rogers and Emma Landers of Gardiner-Caldwell Communications, Macclesfield, UK and was funded by GSK. This systematic review and meta-analysis was funded by GSK. Medical writing support for this manuscript was also funded by GSK.

AUTHOR CONTRIBUTIONS

H.C. made substantial contributions to the design and implementation of the study and analysis of the results, and is accountable for all aspects of the work. He also prepared and reviewed the manuscript, as well as approving and submitting the final version for publication. J.v.d.P. made substantial contributions to the design of the study and is responsible for the accuracy and integrity of the study. He was also closely involved with the development of the manuscript and approved the final version for publication. R.S. reviewed the publications resulting from the searches and identified those to include. He also made substantial contributions to the development of the manuscript, as well as approving the final version for publication. N.B. made substantial contributions to the study design and analysis of the results. He was also closely involved with the development of the manuscript and approved the final version for publication. B.D. planned and carried out the statistical analyses. He also assisted with the development of the manuscript, as well as approving the final version for publication. A.M. made substantial contributions to gathering data and interpretation of results, including the accuracy and integrity of the findings. He was closely involved with the development of the manuscript, including approval the final version for publication. M.T. made substantial contributions to the concept and design of the study, managed the study site in Southampton (UK), analysed data and interpreted the results. He also prepared and revised the manuscript, as well as approving the final version for publication.

COMPETING INTERESTS

H.C. has received sponsorship to carry out studies, together with Board Membership, consultant agreements and honoraria for presentation, from Almirall, AstraZeneca, Boehringer Ingelheim, Chiesi, GSK, Innovata Biomed, Meda, Napp Pharmaceuticals, Mundipharma, NorPharma, Novartis, Orion, Sanofi, Teva, Trudell Medical International, UCB and Zentiva. He has also received research sponsorship from grant awarding bodies (EPSRC and MRC). He is also the owner of Inhalation Consultancy Ltd. J.v.d.P. has received grant and travel funding from GSK. R.S. and N.B. are employed by, and hold stock in, GSK. B.D. has received statistical consulting fees from Bridge Medical, during the conduct of the study and personal fees from Boehringer Ingelheim and from Ipsen outside the submitted work. A.M. is employed by Bridge Medical. M.T. has received personal fees from GSK, Novartis, Aerocrine, Boehringer Ingelheim, Teva, Pfizer, Roche and Sandoz. He has also received non-financial support from GSK, Novartis, AstraZeneca and Boehringer Ingelheim.

REFERENCES

- Al-Jahdali, H. et al. Improper inhaler technique is associated with poor asthma control and frequent emergency department visits. *Allergy Asthma Clin Immunol* **9**, 8 (2013).
- Hesselink, A. E., Penninx, B. W., Wijnhoven, H. A., Kriegsman, D. M. & van Eijk, J. T. Determinants of an incorrect inhalation technique in patients with asthma or COPD. *Scand. J. Prim. Health. Care* **19**, 255–260 (2001).
- GINA report. Global Strategy for asthma management and prevention, updated 2016. Available at: <http://ginasthma.org/2016-gina-report-global-strategy-for-asthma-management-and-prevention/> (accessed August 15, 2016).
- GOLD report. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, updated 2016. Available at: <http://goldcopd.org/global-strategy-diagnosis-management-prevention-copd-2016/> (accessed August 15, 2016).
- Giraud, V. & Roche, N. Misuse of corticosteroid metered-dose inhaler is associated with decreased asthma stability. *Eur. Respir. J.* **19**, 246–251 (2002).
- Lindgren, S., Bake, B. & Larsson, S. Clinical consequences of inadequate inhalation technique in asthma therapy. *Eur. J. Respir. Dis.* **70**, 93–98 (1987).
- Lavorini, F. et al. Effect of incorrect use of dry powder inhalers on management of patients with asthma and COPD. *Respir. Med.* **102**, 593–604 (2008).
- Sanchis, J., Gich, I. & Pedersen, S. Aerosol drug management improvement team (ADMIT). Systematic review of errors in inhaler use: has patient technique improved over time? *Chest* **150**, 394–406 (2016).
- Sarvis, C. A., Tice, A. L. & Plake, K. S. Evaluation of inhaler technique using the aerosol inhalation monitor. *J. Pharm. Pract.* **17**, 275–279 (2004).
- Hilton, S. An audit of inhaler technique among asthma patients of 34 general practitioners. *Br. J. Gen. Pract.* **40**, 505–506 (1990).
- Liard, R. et al. Misuse of pressurized metered dose inhalers by asthmatic patients treated in French private practice. *Rev. Epidemiol. Sante Publ.* **43**, 242–249 (1995).
- Allen, S. C. & Prior, A. What determines whether an elderly patient can use a metered dose inhaler correctly? *Br. J. Dis. Chest* **80**, 45–49 (1986).
- Shrestha, M. et al. Metered-dose inhaler technique of patients in an urban ED: Prevalence of incorrect technique and attempt at education. *Am. J. Emerg. Med.* **14**, 380–384 (1996).
- Molimard, M. et al. Assessment of handling of inhaler devices in real life: An observational study in 3811 patients in primary care. *J. Aerosol. Med. Deposition Clear. Eff. Lung.* **16**, 249–254 (2003).
- Souza, M. L., Meneghini, A. C., Ferraz, E., Vianna, E. O. & Borges, M. C. Knowledge of and technique for using inhalation devices among asthma patients and COPD patients. *J. Bras. Pneumol.* **35**, 824–831 (2009).
- Vargas, O., Martinez, J., Ibanez, M., Pena, C. & Santamaria, M. The use of metered-dose inhalers in hospital environments. *J. Aerosol. Med. Pulm. Drug. Deliv.* **26**, 287–296 (2013).
- Zainudin, B. M. & Sufarlan, A. W. Incorrect use of pressurized metered dose inhaler by asthmatic patients. *Med. J. Malaysia.* **45**, 235–238 (1990).
- Horsley, M. G. & Bailie, G. R. Risk factors for inadequate use of pressurized aerosol inhalers. *J. Clin. Pharm. Ther.* **13**, 139–143 (1988).
- Larsen, J. S., Hahn, M., Ekholm, B. & Wick, K. A. Evaluation of conventional press-and-breathe metered-dose inhaler technique in 501 patients. *J. Asthma.* **31**, 193–199 (1994).
- Erickson, S. R., Horton, A. & Kirking, D. M. Assessing metered-dose inhaler technique: Comparison of observation vs. patient self-report. *J. Asthma.* **35**, 575–583 (1998).
- Plaza, V. & Sanchis, J. Medical personnel and patient skill in the use of metered dose inhalers: A multicentric study. *Respiration.* **65**, 195–198 (1998).
- Al-Hassan, M. I. Assessment of inhaler technique in patients attending a chest hospital in Riyadh City. *Int. J. Pharmacol.* **5**, 232–235 (2009).
- Epstein, S. W., Manning, C. P. R., Ashley, M. J. & Corey, P. N. Survey of the clinical use of pressurized aerosol inhalers. *Can. Med. Assoc. J.* **120**, 813–816 (1979).
- Hashmi, A., Soomro, J. A., Memon, A. & Soomro, T. K. Incorrect inhaler technique compromising quality of life of asthmatic patients. *J. Med.* **13**, 16–21 (2012).
- Arora, P. et al. Evaluating the technique of using inhalation device in COPD and Bronchial Asthma patients. *Respir. Med.* **108**, 992–998 (2014).
- van der Palen, J., Klein, J. J., Kerckhoff, A. H. & van Herwaarden, C. L. Evaluation of the effectiveness of four different inhalers in patients with chronic obstructive pulmonary disease. *Thorax.* **50**, 1183–1187 (1995).
- Khassawneh, B. Y. et al. Handling of inhaler devices in actual pulmonary practice: Metered-dose inhaler versus dry powder inhalers. *Respir. Care.* **53**, 324–328 (2008).
- Melani, A. S. et al. Inhalation technique and variables associated with misuse of conventional metered-dose inhalers and newer dry powder inhalers in experienced adults. *Ann. Allergy. Asthma. Immunol.* **93**, 439–446 (2004).
- Batterink, J., Dahri, K., Aulakh, A. & Rempel, C. Evaluation of the use of inhaled medications by hospital in patients with chronic obstructive pulmonary disease. *Can. J. Hosp. Pharm.* **65**, 111–118 (2012).
- Rootmensen, G. N., Van Keimpema, A. R. J., Jansen, H. M. & de Haan, R. J. Predictors of incorrect inhalation technique in patients with asthma or COPD: A study using a validated videotaped scoring method. *J. Aerosol. Med. Pulm. Drug. Deliv.* **23**, 323–328 (2010).
- Melani, A. S. et al. Inhaler mishandling remains common in real life and is associated with reduced disease control. *Respir. Med.* **105**, 930–938 (2011).
- Ho, S. F., O'Mahony, M. S., Steward, J. A., Brey, P. & Burr, M. L. Inhaler technique in older people in the community. *Age. Ageing.* **33**, 185–188 (2004).
- Giraud, V., Allaert, F. A. & Magnan, A. A prospective observational study of patient training in use of the autohaler (registered trademark) inhaler device: The Sirocco study. *Eur. Rev. Med. Pharmacol. Sci.* **15**, 563–570 (2011).
- Wieshammer, S. & Dreyhaupt, J. Dry powder inhalers: which factors determine the frequency of handling errors? *Respiration.* **75**, 18–25 (2008).
- Chorao, P., Pereira, A. M. & Fonseca, J. A. Inhaler devices in asthma and COPD—An assessment of inhaler technique and patient preferences. *Respir. Med.* **108**, 968–975 (2014).
- Sestini, P. et al. Prescription bias and factors associated with improper use of inhalers. *J. Aerosol. Med. Deposition. Clear. Eff. Lung.* **19**, 127–136 (2006).
- Brennan, V. K., Osman, L. M., Graham, H., Critchlow, A. & Everard, M. L. True device compliance: The need to consider both competence and contrivance. *Respir. Med.* **99**, 97–102 (2005).
- Van Beerendonk, I., Mesters, I., Mudde, A. M. & Tan, T. D. Assessment of the inhalation technique in outpatients with asthma or chronic obstructive pulmonary disease using a metered-dose inhaler or dry powder device. *J. Asthma.* **35**, 273–279 (1998).
- van der Palen, J., Klein, J. J., van Herwaarden, C. I. A., Zielhuis, G. A. & Seydel, E. R. Multiple inhalers confuse asthma patients. *Eur. Respir. J.* **14**, 1034–1037 (1999).
- Serra-Batllés, J., Plaza, V., Badiola, C. & Morejón, E. Patient perception and acceptability of multidose dry powder inhalers: A randomized crossover comparison of Diskus/Accuhaler with Turbuhaler. *J. Aerosol. Med. Deposition. Clear. Eff. Lung.* **15**, 59–64 (2002). Inhalation Devices Study Group.
- De Blaquiere, P., Christensen, D. B., Carter, W. B. & Martin, T. R. Use and misuse of metered-dose inhalers by patients with chronic lung disease: A controlled, randomized trial of two instruction methods. *Am. Rev. Respir. Dis.* **140**, 910–916 (1989).
- Goodman, D. E. et al. The influence of age, diagnosis and gender on proper use of metered-dose inhalers. *Am. J. Respir. Crit. Care. Med.* **150**, 1256–1261 (1994).
- Dalcin Pde, T. et al. Factors related to the incorrect use of inhalers by asthma patients. *J. Bras. Pneumol.* **40**, 13–20 (2014).
- Dahl, R., Backer, V., Ollgaard, B., Gerken, F. & Kesten, S. Assessment of patient performance of the HandiHaler (registered trademark) compared with the metered dose inhaler four weeks after instruction. *Respir. Med.* **97**, 1126–1133 (2003).
- van der Palen, J., Eijsvogel, M. E., Kuipers, B. F., Schipper, M. & Vermue, N. A. Comparison of the Diskus[®] Inhaler and the Handihaler[®] regarding preference and ease of use. *J. Aerosol. Med.* **20**, 38–44 (2007).
- Schulte, M. et al. Handling of and preferences for available dry powder inhaler systems by patients with asthma and COPD. *J. Aerosol. Med. Pulm. Drug. Deliv.* **21**, 321–328 (2008).
- Liam, C.-K., Lim, K.-H. & Wong, C. M. M. Acceptance of the Accuhaler, a multidose powder inhaler, among asthmatic patients: A comparison with the pressurized metered-dose inhaler. *Asian. Pac. J. Allergy. Immunol.* **18**, 135–140 (2000).
- Basheti, I. A. et al. User error with diskus and turbuhaler by asthma patients and pharmacists in Jordan and Australia. *Respir. Care.* **56**, 1916–1923 (2011).
- Epstein, S., Maidenberg, A., Hallett, D., Khan, K. & Chapman, K. R. Patient handling of a dry-powder inhaler in clinical practice. *Chest* **120**, 1480–1484 (2001).
- Ronmark, E. et al. Correct use of three powder inhalers: Comparison between diskus, turbuhaler, and easyhaler. *J. Asthma.* **42**, 173–178 (2005).
- Long Alliantie Nederland, LAN. <http://inhalatorgebruik.nl/nl/zorgverleners/inhalatieprotocolen> (accessed January 5, 2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *npj Primary Care Respiratory Medicine* website (doi:10.1038/s41533-017-0016-z).