



OPEN

A data-driven typology of asthma medication adherence using cluster analysis

Holly Tibble^{1,2}✉, Amy Chan^{3,4}, Edwin A. Mitchell⁵, Elsie Horne^{1,2}, Dimitrios Doudehis^{1,6}, Rob Horne^{2,4}, Mehrdad A. Mizani^{1,2}, Aziz Sheikh^{1,2,7} & Athanasios Tsanas^{1,2}

Asthma preventer medication non-adherence is strongly associated with poor asthma control. One-dimensional measures of adherence may ignore clinically important patterns of medication-taking behavior. We sought to construct a data-driven multi-dimensional typology of medication non-adherence in children with asthma. We analyzed data from an intervention study of electronic inhaler monitoring devices, comprising 211 patients yielding 35,161 person-days of data. Five adherence measures were extracted: the percentage of doses taken, the percentage of days on which zero doses were taken, the percentage of days on which both doses were taken, the number of treatment intermissions per 100 study days, and the duration of treatment intermissions per 100 study days. We applied principal component analysis on the measures and subsequently applied k-means to determine cluster membership. Decision trees identified the measure that could predict cluster assignment with the highest accuracy, increasing interpretability and increasing clinical utility. We demonstrate the use of adherence measures towards a three-group categorization of medication non-adherence, which succinctly describes the diversity of patient medication taking patterns in asthma. The percentage of prescribed doses taken during the study contributed to the prediction of cluster assignment most accurately (84% in out-of-sample data).

Abbreviations

CART	Classification and regression tree
EMD	Electronic monitoring device
ICS	Inhaled corticosteroids
IQR	Interquartile range
PC	Principal component
PCA	Principal component analysis
RCT	Randomized controlled trial
US	United States

Asthma is a long-term condition that can be effectively managed in the majority of individuals through regular use of inhaled corticosteroids (ICS)^{1–3}. *Medication non-adherence*, that is not taking a medication as prescribed, is however very common and is a major factor associated with poor asthma control, leading to increased likelihood of asthma attacks^{4–8}. Sub-optimal adherence is a costly problem due to medication waste and unnecessary additional healthcare utilization to treat uncontrolled symptoms that may have arisen from non-adherence^{9–14}.

Numerous studies have used targeted interventions to improve medication adherence in patients with asthma, with the aim of improving clinical outcomes (such as reducing the risk of asthma symptoms or attacks). While

¹Usher Institute, Edinburgh Medical School, College of Medicine and Veterinary Medicine, University of Edinburgh, Doorway 1, Old Medical School, Teviot Place, Edinburgh EH8 9AG, UK. ²Asthma UK Centre for Applied Research, Usher Institute, Centre for Medical Informatics, University of Edinburgh, Edinburgh, UK. ³School of Pharmacy, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand. ⁴Centre for Behavioural Medicine, Department for Practice and Policy, UCL School of Pharmacy, University College London, London, UK. ⁵Department of Paediatrics: Child and Youth Health, Faculty of Medicine and Health Sciences, University of Auckland, Auckland, New Zealand. ⁶BHF Centre for Cardiovascular Sciences, University of Edinburgh, Edinburgh, UK. ⁷Health Data Research UK, London, UK. ✉email: Holly.Tibble@ed.ac.uk

some studies have shown positive changes in adherence behavior, they are usually short in duration, making their true clinical and cost-effectiveness hard to assess^{15–25}.

Electronic monitoring devices (EMDs) enable the real-time tracking of inhaler use, by means of a small electronic chip fitted onto an inhaler, which records the date and time of each dose taken. EMDs are commonly used in research studies aimed at measuring adherence to inhaler asthma therapy²⁶, as they can track the exact time and frequency of inhaler actuations. EMDs can thus measure medication use more accurately than either self-report or indirect adherence measurement methods such as using prescribing or dispensing records, which makes their data a valuable resource for investigating patterns of adherence behaviors.

There are many ways of measuring medication non-adherence and the correlation between these measures within the same data is heavily context dependant^{27,28}. There is currently no consensus on what should be considered the gold standard^{27–30}. Furthermore, despite the multitude of medication non-adherence measures³¹, many intervention studies describe medication adherence using only a single average measure over time (such as reporting the percentage of days on which the prescribed doses of ICS were taken^{32–36}), without supporting rationale for their choice.

Previous studies of adherence using EMDs in asthma have most commonly used the proportion of prescribed doses that were taken^{34,37–40}, or the percentage of days on which adherence was deemed acceptable by the authors, according to timing of doses^{38,41}.

Adherence is often reported categorically by dichotomizing individuals into “good” and “poor” adherers, according to a binary threshold, commonly 80%^{40,42–50}. This threshold, based on a hypothesised level of adherence required to maintain therapeutic effect, which may be considered arbitrary, as any such level would be condition and drug-specific⁵¹ and rely on consistency in the space between doses over time^{52–54}. Very few studies have attempted a more detailed and meaningful dissection, such as categorising based on trends over time^{55,56}, or by assessing multiple components of adherence^{7,57,58}.

This study aimed to investigate the underlying patterns of medication non-adherence in children with asthma, and to identify multidimensional behavior subgroups.

Methods

Data sources. We undertook a secondary analysis of data from a previously published randomized controlled trial (RCT)³⁷. Chan et al.³⁷ investigated the effect of an EMD inhaler with an audio-visual reminder function (timed alarm) on ICS adherence. The study enrolled children who presented to the emergency department for acute asthma symptoms and were diagnosed with asthma by the emergency department medical team. The diagnosis process would usually include a clinical diagnosis based on the presentation and medical assessment, supported by pulmonary function tests.

There were 220 children enrolled, aged 6–15 years, of whom 110 were randomized to the control group, and 110 to the intervention group. Both groups received the same EMD; however, in the intervention group, the EMD had the audio-visual function enabled, whilst the control group had this function disabled. The device was provided to study participants at the baseline appointment, and data from each device was collected at each of the three follow-up appointments (approximately week 8, week 16, and week 24).

The original study inclusion criteria were: children between 6 and 15 years, with physician-diagnosed asthma, presenting to the emergency department with an asthma attack, and on twice-daily ICS treatment (with or without additional asthma treatments). Diagnosis of a chronic lung disease other than asthma, or congenital heart disease, led to exclusion from participating in the study, as well as living outside of the Auckland, New Zealand catchment area.

Measures of adherence. In chronic disease management, it is common for patients to take *intermissions* of treatment (periods of consecutively missed doses), followed by *re-initiations* (returning to treatment following an intermission). As such, it is possible to categorise adherence measures into the subdomains of *implementation* (dose-taking within a day³¹) and *persistence* (the incidence and duration of treatment intermissions lasting at least five days, as discussed by Andrade et al.⁵⁹).

Twice daily asthma medications are typically taken in the morning and evening, to ensure consistent bio-availability throughout a 24-h period. As such, we refer herein to a ‘dose’ as one or more actuations (puffs) of the inhaler during a 12-h period (midnight to midday, and midday to midnight). Each individual in this study, by virtue of the inclusion criteria, was on a twice daily regimen, and thus had two potential doses in a day (before and after midday). Herein we define five measures of adherence used in this analysis, calculated at the person-level, with denominators excluding periods in which the electronic inhaler device was malfunctioning. Their algorithmic definitions are as follows:

The percentage of doses that were taken:

$$\frac{\text{Doses taken}}{2 \times \text{Days with device}} \times 100 \quad (\text{A})$$

The percentage of days on which zero doses were taken:

$$\frac{\text{Days on which zero doses were actuated}}{\text{Days with device}} \times 100 \quad (\text{B})$$

The percentage of days on which both doses were taken:

$$\frac{\text{Days on which doses were actuated in both the AM and PM}}{\text{Days with device}} \times 100 \quad (\text{C})$$

The number of treatment intermissions, per 100 days of follow-up:

$$\frac{\text{Number of treatment intermissions}}{\text{Days with device}} \times 100 \quad (\text{D})$$

The duration of time spent on treatment intermissions, per 100 days of follow-up:

$$\frac{\text{Total duration of time spent on treatment intermission}}{\text{Days with device}} \times 100 \quad (\text{E})$$

Data analysis. Using principal component analysis (PCA), we extracted the *latent variables* of the adherence measures. Latent variables (also known as *principal components*) are linear functions of the original measures, which identify the underlying structure of the measures and avoid redundant information⁶⁰. We identified the minimum number of principal components required to explain at least 95% of the variance between the measures.

Principal components can be influenced by varying magnitudes of scale (e.g. range 0–1 versus range 0–1,000) in the input features. As a benchmark, we used only *zero-centering* (subtracting the mean from each value) to pre-process the adherence measures prior to applying PCA. We also compared the computed principal components generated from the adherence measures when the measures underwent different pre-processing before applying PCA: (1) after scaling each variable to *unit variance* (subtracting the mean from each value, and dividing by the variance), and (2) using *min–max scaling* (each variable was scaled to lie between 0 and 1).

We used k-means clustering with 25 random initiations, searching through for the optimal number *k* of clusters (we searched for *k* = 1, 2, ..., 15). We visually assessed the within groups sum of squares using a *scree plot*⁶¹, to determine the optimal number of clusters. For each method of scaling/centering applied to the data in the principal component computation, we performed clustering on bootstrapped samples of the principal components, repeating the process 1,000 times to assess the stability of the resulting clusters.

The optimal scaling method was selected as the one that led to the most stable clusters, formally assessed using measures derived from the Jaccard similarity coefficient.

The Jaccard coefficient of cluster solution similarity⁶² is the mean Jaccard similarity between (1) the clusters constructed from the full sample, and (2) the corresponding cluster constructed from bootstrapped samples. The cluster in the bootstrap sample corresponding to the original cluster is determined as the one with the highest Jaccard similarity to the original cluster.

In this process, a bootstrap sample cluster is considered *dissolved* when the Jaccard similarity between the full data cluster and the associated cluster from a bootstrap sample is less than 0.5⁶³. As such, we can also measure the cluster dissolution rate, the proportion of bootstrap samples for which the cluster dissolved.

Finally, we trained standard classification and regression trees (CARTs)⁶⁰, using each of the measures described above to estimate the adherence groups, using the computed cluster labels as ground truth. The trees were constructed using the Breiman et al. algorithm⁶⁴ with the following parameters:

- At least 20 patients must be at a node for a split to be considered,
- There must be at least 6 patients at any terminal node,
- The maximum tree height is 2 nodes,

tenfold cross-validation was conducted to estimate the optimal *complexity parameter* (threshold improvement in model fit required for a split to be retained) for use in the tree pruning in order to minimize the cross-validated error rate. The tree is then pruned according to this estimated complexity parameter.

The tree was built on 70% of the data (randomly sampled). A partition of 15% was used to generate a confusion matrix for each measure; mapping the combinations of cluster labels and tree labels, to quantify the quality of the estimates produced by the decision tree. The best measure was selected by highest accuracy was identified, as follows:

$$\text{Accuracy} = \frac{\text{Number of correctly classified patients}}{\text{Number of patients}}$$

We then used the remaining 15% of the data (the test set) to report the final accuracy.

All analyses were conducted in R version 3.5.1, and scripts for data cleaning, setup and analysis will be made available on GitHub, a platform which host open-source projects (<https://github.com/hollytibble/Electronic-Monitoring-Device-Adherence-Typology>).

Ethics. Ethics approval was granted by the New Zealand Northern Y Regional Ethics Committee (NTY/08/12/116) and District Health Boards. Verbal and written informed consent was obtained at data collection from the parent or guardian of all participants. Analysis and reporting in this study was performed in accordance with the relevant guidelines and regulations.

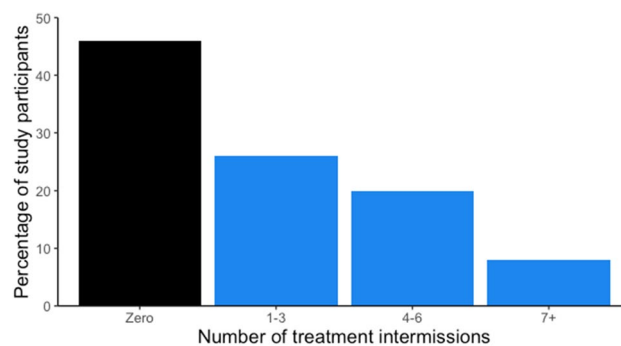


Figure 1. Histogram of the number of treatment intermissions (continuous abstinence of 5 days or longer) per person during study period.

	PC1	PC2	PC3	PC4
(A) Percentage of doses taken	0.98	-0.18	-0.03	0.02
(B) Percentage of days on which zero doses were taken	-0.99	-0.04	0.10	-0.09
(C) Percentage of days on which both doses were taken	0.93	-0.36	0.04	-0.05
(D) Number of treatment intermissions per 100 study days	-0.94	-0.28	0.19	0.07
(E) Duration of treatment intermissions per 100 study days	-0.93	-0.23	-0.29	0.00
Percentage of variance explained	91%	6%	<1%	<1%

Table 1. Principal components to identify the latent variable structure of (unit-scaled) medication non-adherence.

Results

Data cleaning and missing data. A total of 220 children were enrolled into the study. There were nine participants (seven from the control group and two from the intervention group) who did not complete the study. Reasons for drop-out were loss to follow-up ($n=5$), missing data from one or more follow-up appointment ($n=2$), non-compliance with protocol ($n=1$) and study withdrawal ($n=1$). This left a total of 211 participants with data from all study follow-up periods for inclusion in this analysis, yielding 36,902 person-days of data.

There were 1,741 person-days of missing data, out of a total 36,902 person-days (5% missing). This was due to battery failure, indicated in the device log, rather than simply individuals missing doses. 55 (26%) children in the study sample had at least one day with missing data, with a median 22 days missing (interquartile range (IQR): 10–48, range 2–122) from a median 185 days of follow-up (IQR: 176–190, range: 155–210). These periods of missing data were removed, resulting in 35,161 person-days for analysis.

The peak hours for actuations were 7–10 AM (96% of actuations between midnight and midday) and 7–10 AM (93% of actuations between midday and midnight).

Study population. The study population consisted of 108 boys and 103 girls, aged between 6 and 15 (median 8) years. The most common ethnicity was White European origin (38%; eTable 1).

Adherence. In the control group, 36% of prescribed doses were taken, compared to 80% in the intervention group. 54% of the overall cohort ($n=114$) had at least one treatment intermission (Fig. 1), with the first starting a median of 42 days after the baseline visit (IQR 14–95 days). Individual's first treatment intermissions had a median duration of eight days (IQR 6–14 days).

The difference between the intervention arms was marked, with 27% ($n=29$) of the intervention group participants having at least one treatment intermission, compared to 83% ($n=85$) in the control group.

Exploratory cluster analysis. The correlations between the adherence measures used herein have previously been reported⁶⁵. For each scaling method, the scree plot consistently found three as the most appropriate number of clusters. Computing the principal components from the unit-variance scaled measures gave the most robust solution (eTable 2). Table 1 shows the variable loadings for the first four principal components.

Determining clinically meaningful subgroups. Statistics to summarize the distribution of the measures by cluster are presented in eTable 3. Cluster 1 had the poorest adherence—taking on average (between individuals median) 16% of their prescribed doses and approximately 4.0 treatment intermissions of five or more days in length for every 100 days of follow-up. In Cluster 3, only 7 individual (11%) took any treatment

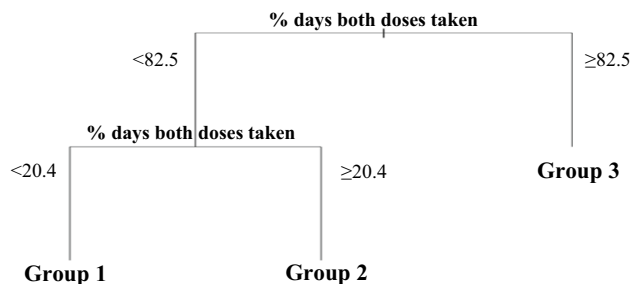


Figure 2. Decision tree for estimating exploratory adherence clusters.

n = 32	Exploratory clusters		
	C1	C2	C3
CART groups			
G1	4	2	0
G2	1	18	0
G3	0	0	5

Table 2. Confusion matrix using the automatically determined cluster labels (used as 'ground truth') and CART estimates in the testing set.

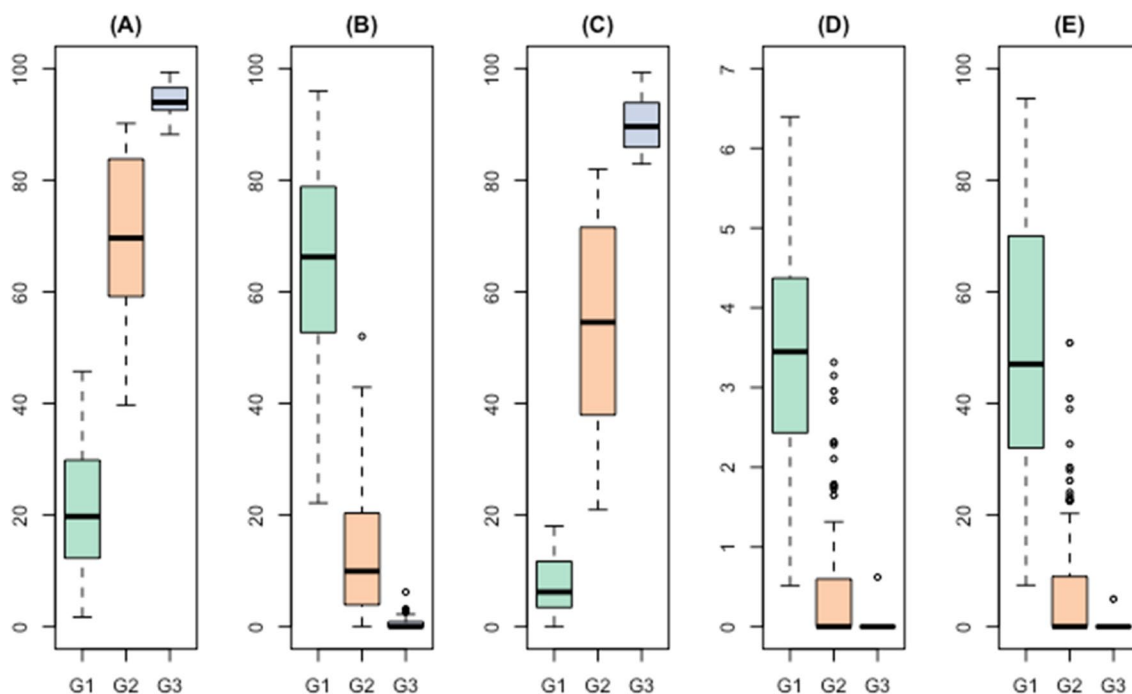


Figure 3. Boxplots of the five adherence measures by tree-derived groups. (A) Percentage of doses taken, (B) Percentage of days zero doses taken, (C) Percentage of days both doses taken, (D) Number of treatment intermissions per 100 study days, (E) Duration of treatment intermissions per 100 study days.

intermissions, and on average they took 91% of their doses. For each measure, Cluster 2 was somewhere between the first and third clusters.

We constructed a decision tree for each adherence measure in isolation, using the cluster labels as the response variable, in a standard supervised learning set-up framework. We used a random 70% sample of the data as a training set and the remaining 30% as a testing set. The percentage of doses taken yielded the highest accuracy (i.e. the group using the decision tree matching the group computed using the exploratory clustering approach) at 88%. The resulting decision tree (Fig. 2) had an accuracy of 84% in the test set (Table 2). Figure 3A–E present the boxplots of the measures by group.

Discussion

In this study, we identified three exploratory clusters of adherence, which were found to be very stable across the 1,000 bootstrapped repetitions to assess cluster stability. These clusters, representing poor, moderate, and good medication adherence, were best approximated in a decision tree using the measure of the percentage of prescribed doses that were taken during the study (84% accuracy in data unseen to the training model).

Due to the limited follow-up time, we were required to aggregate adherence measures over the duration of the study. It is likely that with a larger sample and a longer follow-up we may have observed transitioning between states of behavior, perhaps triggered by clinical events such as an asthma attack. This study used data from 211 children, and hence findings should be interpreted tentatively: it is possible that having more samples would enable finer stratification and exploration of additional clusters.

Some individuals with asthma are instructed to take their medication 'as needed', as they are able to maintain adequate asthma control despite taking medicine inconsistently. We did not remove such children from the population, as their adherence patterns remain of interest. We were additionally unable to stratify the population accordingly due to the small sample size. In a larger study, it might have been that different groups were detected according to the strictness of their regimen, or that observed groups occurred in different frequencies.

The proportion of days on which both daily doses were taken was used as one of the measures of adherence. This measure was calculated assuming that one dose (of at least one actuation) was taken between midnight and midday, and one between midday and midnight. For the most part, this meant one dose in the morning, and one in the evening, as is usually recommended. However, if a dose was taken at 11.30 in the morning, followed by another at 12.30 in the afternoon, this would be recorded as having taken both the morning and evening doses, despite their being insufficient time between the doses for optimal distribution throughout the day. In this cohort, over 96% of morning and 93% of evening doses were taken within a 3-h window in each period (between 7:00–10:00 o'clock in the morning window, and between 19:00–22:00 in the afternoon window). As such, this is unlikely to have been a common concern in the context of this study.

Finally, this study was undertaken in a narrow cohort with narrowly defined inclusion criteria i.e. children on a twice daily regimen for clinician-diagnosed asthma, with a recent history of emergency care presentation for an asthma attack. The generalizability of these findings to all children diagnosed with asthma, (or adults diagnosed with asthma) would therefore need to be further studied.

Future work. There is a need to study adherence patterns in adults with asthma, as well as in children (and adults) in primary care setting rather than only using a cohort of those with asthma related emergency department presentations. As technology advances and the cost of manufacturing EMDs decreases, we envisage an increase in their adoption in clinical care^{66,67}, which may mean that future large scale studies can be conducted on routine practice datasets.

There are no studies which have attempted data driven categorisation, as far as we are aware. A recent study by Allemann et al. generated simulated pharmacy refill records according to six proposed adherence subtypes and attempted to detect them using data driven trajectory-based models⁶⁸. Trajectory-based models offer little scope for the integration of domain-specific knowledge. As the authors note, this means that identifying clusters which are clinically interpretable and meaningful according to their temporal patterns is challenging. While electronic monitoring devices are considered the gold standard for adherence measurement in asthma, adapting our methods to refill data may result in the identification of some of the clusters described by the authors. Furthermore, the application of such trajectory-based models in a (larger) electronic monitoring device dataset could potentially uncover some previously unidentified subgroups.

Conclusions

Medication taking behavior in children with asthma is highly heterogeneous, and there is no consistent pattern of a patient with poor adherence. Through analysis of multiple adherence measures, we identified subgroups of asthma patients according to data-driven thresholds. Our findings suggest that the simple measurement of percentage of doses taken is sufficient in this population, and that combinations measures of non-adherence are unlikely to be of further clinical value than any of these measures in isolation.

Received: 11 October 2019; Accepted: 20 August 2020

Published online: 14 September 2020

References

1. Barnes, P. J. Efficacy of inhaled corticosteroids in asthma. *J. Allergy Clin. Immunol.* **102**, 531–538 (1998).
2. Barnes, P. J. & Pedersen, S. Efficacy and safety of inhaled corticosteroids in asthma. *Am. Rev. Respir. Dis.* **148**, S1–S26 (1993).
3. Suissa, S., Ernst, P., Benayoun, S., Baltzan, M. & Cai, B. Low-dose inhaled corticosteroids and the prevention of death from asthma. *N. Engl. J. Med.* **343**, 332–336 (2000).
4. Urquhart, J. Pharmacoeconomic consequences of variable patient compliance with prescribed drug regimens. *Pharmacoeconomics* **15**, 217–228 (1999).
5. Papi, A. *et al.* Relationship of inhaled corticosteroid adherence to asthma exacerbations in patients with moderate-to-severe asthma. *J. Allergy Clin. Immunol. Pract.* <https://doi.org/10.1016/j.jaip.2018.03.008> (2018).
6. Boslev, C., Md, B., Suppli, C. & Dmsc, U. M. Asthma and adherence to inhaled corticosteroids: Current status and future perspectives. *Respir. Care* **60**, 455–468 (2015).
7. Engelkes, M., Janssens, H. M., De Jongste, J. C., Sturkenboom, M. C. J. M. & Verhamme, K. M. C. Medication adherence and the risk of severe asthma exacerbations: A systematic review. *Eur. Respir. J.* **45**, 396–407 (2015).
8. Guedes, A. *et al.* Risk factors for death in patients with severe asthma. *J. Bras. Pneumol.* **40**, 364–372 (2014).

9. Van Boven, J. F. M. *et al.* Clinical and economic impact of non-adherence in COPD: A systematic review. *Respir. Med.* **108**, 103–113 (2014).
10. Hommel, K. A. *et al.* Longitudinal patterns of medication nonadherence and associated health care costs. *Inflamm. Bowel Dis.* **23**, 1577–1583 (2017).
11. Iuga, A. O. & McGuire, M. J. Adherence and health care costs. *Risk Manag. Healthc. Policy* **7**, 35–44 (2014).
12. Cutler, R. L., Fernandez-Llimos, F., Frommer, M., Benrimoj, C. & Garcia-Cardenas, V. Economic impact of medication non-adherence by disease groups: A systematic review. *BMJ Open* **8**, e016982 (2018).
13. Ho, S. C., Chong, H. Y., Chaiyakunapruk, N., Tangiisuran, B. & Jacob, S. A. Clinical and economic impact of non-adherence to antidepressants in major depressive disorder: A systematic review. *J. Affect. Disord. J.* **193**, 1–10 (2016).
14. Wu, A. C. *et al.* Primary adherence to controller medications for asthma is poor. *Ann. Am. Thorac. Soc.* **12**, 161–166 (2015).
15. Fireman, P., Friday, G. A., Gira, C., Vierthaler, W. A. & Michaels, L. Teaching self-management skills to asthmatic children and their parents in an ambulatory care setting. *Pediatrics* **68**, 341–348 (1981).
16. Lewis, C. E. *et al.* A randomized trial of A.C.T. (Asthma Care Training) for kids. *Pediatrics* **74**, 478–486 (1984).
17. Zullig, L. L., Deschodt, M., Liska, J., Bosworth, H. B. & Geest, S. D. Moving from the trial to the real world: Improving medication adherence using insights of implementation science. *Annu. Rev. Pharmacol. Toxicol.* <https://doi.org/10.1146/annurev-pharmtox-010818-021348> (2019).
18. McNabb, W. L., Wilson-Pessano, S. R., Hughes, G. W. & Scamagas, P. Self-management education of children with asthma: AIR WISE. *Public Health* **75**, 1219–1220 (1985).
19. Weingarten, M. A., Goldberg, J., Teperberg, Y., Harrison, N. & Oded, A. A pilot study of the multidisciplinary management of childhood asthma in a family practice. *J. Asthma* **22**, 261–265 (1985).
20. Shields, M. C., Griffin, K. W. & McNabb, W. L. The effect of a patient education program on emergency room use for inner-city children with asthma. *Am. J. Public Health* **80**, 36–38 (1990).
21. Hughes, D. M., McLeod, M., Garner, B. & Goldbloom, R. B. Controlled trial of a home and ambulatory program for asthmatic children. *Pediatrics* **87**, 54–60 (1991).
22. Elliott, R. A., Barber, N. & Horne, R. Cost-effectiveness of adherence-enhancing interventions: A quality assessment of the evidence. *Ann. Pharmacother.* **39**, 508–515 (2005).
23. Weinstein, A. G. *et al.* A pilot study of the effect of an educational web application on asthma control and medication adherence. *J. Allergy Clin. Immunol. Pract.* **7**, 1497–1506 (2019).
24. Riley, I. L. *et al.* A systematic review of patient- and family-level inhaled corticosteroid adherence interventions in Black/African Americans. *J. Allergy Clin. Immunol. Pract.* **7**, 1184–1193.e3 (2019).
25. Normansell, R., Kew, K. M. & Mathioudakis, A. G. Interventions to improve inhaler technique for people with asthma. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD012286.pub2> (2017).
26. Chan, A. H. Y. *et al.* Adherence monitoring and e-health: How clinicians and researchers can use technology to promote inhaler adherence for asthma. *J. Allergy Clin. Immunol. Pract.* **1**, 446–454 (2013).
27. Hess, L. M., Raebel, M. A., Conner, D. A. & Malone, D. C. Measurement of adherence in pharmacy administrative databases: A proposal for standard definitions and preferred measures. *Ann. Pharmacother.* **40**, 1280–1288 (2006).
28. Sajatovic, M., Velligan, D. I., Weiden, P. J., Valenstein, M. A. & Ogedegbe, G. Measurement of psychiatric treatment adherence. *J. Psychosom. Res.* **69**, 591–599 (2010).
29. Lehmann, A. *et al.* Assessing medication adherence: Options to consider. *Int. J. Clin. Pharm.* **36**, 55–69 (2014).
30. Williams, A. B., Amico, K. R., Bova, C. & Womack, J. A. A proposal for quality standards for measuring medication adherence in research. *AIDS Behav.* **17**, 284–297 (2013).
31. Vrijens, B. *et al.* A new taxonomy for describing and defining adherence to medications. *Br. J. Clin. Pharmacol.* **73**, 691–705 (2012).
32. Chatkin, J. M. *et al.* Impact of a low-cost and simple intervention in enhancing treatment adherence in a Brazilian asthma sample. *J. Asthma* **43A**, 263–266 (2006).
33. Bender, B. G. *et al.* Test of an interactive voice response intervention to improve adherence to controller medications in adults with asthma. *J. Am. Board Fam. Med.* **23**, 159–165 (2010).
34. Charles, T. *et al.* An audiovisual reminder function improves adherence with inhaled corticosteroid therapy in asthma. *J. Allergy Clin. Immunol.* **119**, 811–816 (2007).
35. Ducharme, F. M. *et al.* Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control. *Am. J. Respir. Crit. Care Med.* **183**, 195–203 (2011).
36. Foster, J. M. *et al.* Inhaler reminders improve adherence with controller treatment in primary care patients with asthma. *J. Allergy Clin. Immunol.* **134**, 1260–1268.e3 (2014).
37. Chan, A. H. Y. *et al.* The effect of an electronic monitoring device with audiovisual reminder function on adherence to inhaled corticosteroids and school attendance in children with asthma: A randomised controlled trial. *Lancet Respir. Med.* **3**, 210–219 (2015).
38. Chung, K. F. & Naya, I. Compliance with an oral asthma medication: A pilot study using an electronic monitoring device. *Respir. Med.* **94**, 852–858 (2000).
39. Jentzsch, N. S., Camargos, P. A. M., Colosimo, E. A. & Bousquet, J. Monitoring adherence to beclomethasone in asthmatic children and adolescents through four different methods. *Allergy Eur. J. Allergy Clin. Immunol.* **64**, 1458–1462 (2009).
40. Garcia-Marcos, P. W., Brand, P. L. P., Kaptein, A. A. & Klook, T. Is the MARS questionnaire a reliable measure of medication adherence in childhood asthma?. *J. Asthma* **53**, 1085–1089 (2016).
41. Nides, M. A. *et al.* Improving inhaler adherence in a clinical trial through the use of the nebulizer chronolog*. *Chest* **104**, 501–507 (1993).
42. Haupt, D. & Nilsson, J. L. G. Asthma medication—Persistence with adrenergics, steroids and combination products over a 5-year period. *J. Clin. Pharm. Ther.* **34**, 507–513 (2009).
43. Salazar, G., Tarwala, G. & Reznik, M. School-based supervised therapy programs to improve asthma outcomes: Current perspectives. *J. Asthma Allergy* **11**, 205–215 (2018).
44. Krishnan, J. A. *et al.* Adherence to inhaled corticosteroids: An ancillary study of the Childhood Asthma Management Program clinical trial. *J. Allergy Clin. Immunol.* **129**, 112–118 (2012).
45. Creer, T. L. & Levstek, D. Adherence to asthma regimens. In *Handbook of Health Behavior Research* (ed. Gochman, D. S.) (Springer, Boston, 1997). https://doi.org/10.1007/978-1-4899-1760-7_7.
46. George, M. & Bender, B. New insights to improve treatment adherence in asthma and COPD. *Patient Prefer. Adherence* <https://doi.org/10.2147/PPA.S209532> (2019).
47. Murphy, A. C. *et al.* The relationship between clinical outcomes and medication adherence in difficult-to-control asthma. *Thorax* **67**, 751–753 (2012).
48. Feehan, M. *et al.* Adherence to controller asthma medications: 6-month prevalence across a US community pharmacy chain. *J. Clin. Pharm. Ther.* **40**, 590–593 (2015).
49. Ismaila, A. *et al.* Impact of adherence to treatment with fluticasone propionate/salmeterol in asthma patients. *Curr. Med. Res. Opin.* **30**, 1417–1425 (2014).
50. Arcoleo, K. J. *et al.* Longitudinal patterns of Mexican and Puerto Rican children's asthma controller medication adherence and acute healthcare use. *Ann. Am. Thorac. Soc.* **16**, 715–723 (2019).

51. Stauffer, M. E., Hutson, P., Kaufman, A. S. & Morrison, A. The adherence rate threshold is drug specific. *Drugs R&D* **17**, 645–653 (2017).
52. Morrison, A., Stauffer, M. E. & Kaufman, A. S. Relationship between adherence rate threshold and drug ‘forgiveness’. *Clin. Pharmacokinet.* **56**, 1435–1440 (2017).
53. Urquhart, J. Role of patient compliance in clinical pharmacokinetics: A review of recent research. *Clin. Pharmacokinet.* **27**, 202–215 (1994).
54. Boissel, J. P. & Nony, P. Using pharmacokinetic-pharmacodynamic relationships to predict the effect of poor compliance. *Clin. Pharmacokinet.* **41**, 1–6 (2002).
55. Knafelz, G. J., Schoenthaler, A. & Ogedegbe, G. Secondary analysis of electronically monitored medication adherence data for a cohort of hypertensive African-Americans. *Patient Prefer. Adherence* **6**, 207–219 (2012).
56. Dillon, P., Stewart, D., Smith, S. M., Gallagher, P. & Ainnie Cousins, G. Group-based trajectory models: Assessing adherence to antihypertensive medication in older adults in a community pharmacy setting. *Clin. Pharmacol. Ther.* **103**, 1052–1060 (2018).
57. Holvast, F. *et al.* Non-adherence to cardiovascular drugs in older patients with depression: A population-based cohort study. *Int. J. Cardiol.* **274**, 366–371 (2019).
58. Holvast, F. *et al.* Non-adherence to antidepressants among older patients with depression: A longitudinal cohort study in primary care. *Fam. Pract.* <https://doi.org/10.1093/fampra/cmz106> (2018).
59. Andrade, S. E., Kahler, K. H., Frech, F. & Chan, K. A. Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiol. Drug Saf.* **15**, 565–574 (2006).
60. Hastie, T., Tibshirani, R. & Friedman, J. *Elements of Statistical Learning. Springer Series in Statistics* 2nd edn. (Springer, New York, 2009).
61. Cattell, R. B. The scree test for the number of factors. *Multivar. Behav. Res.* **1**, 245–276 (1966).
62. Hennig, C. Cluster-wise assessment of cluster stability. *Comput. Stat. Data Anal.* <https://doi.org/10.1016/j.csda.2006.11.025> (2007).
63. Hennig, C. Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. *J. Multivar. Anal.* **99**, 1154–1176 (2008).
64. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and Regression Trees. Classification and Regression Trees* (Routledge, London, 1984). <https://doi.org/10.1201/9781315139470>.
65. Tibble, H. *et al.* Heterogeneity in asthma medication adherence measurement. In *BIBE 2019; International Conference on Biological Information and Biomedical Engineering*.
66. Chan, A. H. Y., Harrison, J., Black, P. N., Mitchell, E. A. & Foster, J. M. Using electronic monitoring devices to measure inhaler adherence: A practical guide for clinicians. *J. Allergy Clin. Immunol. Pract.* **3**, 335–349.e5 (2015).
67. Howard, S., Lang, A., Sharples, S. & Shaw, D. What are the pros and cons of electronically monitoring inhaler use in asthma? A multistakeholder perspective. *BMJ Open Respir. Res.* **3**, 1–10 (2016).
68. Allemann, S. S., Dediu, D. & Dima, A. L. Beyond adherence thresholds: A simulation study of the optimal classification of longitudinal adherence trajectories from medication refill histories. *Front. Pharmacol.* **10**, 1–13 (2019).

Author contributions

H.T., E.H., D.D. and A.T. conceived and planned the analysis. H.T. wrote the first draft, with contributions from E.H., A.C., E.A.M., D.D., R.H., M.A.M., A.S., and A.T. All authors approved the final version and jointly take responsibility for the decision to submit this manuscript to be considered for publication.

Funding

HT is supported by College of Medicine and Veterinary Medicine PhD (eHERC/Farr Institute) Studentships from The University of Edinburgh. EH is supported by a Medical Research Council PhD Studentship (eHERC/Farr). DD is supported by a Medical Research Council PhD Studentship (DTP Precision Medicine). MAM’s Newton International Fellowship is awarded by the Academy of Medical Sciences and Newton Fund. This work is carried out with the support of the Asthma UK Centre for Applied Research [AUK-AC-2012-01] and Breathe - The Health Data Research Hub for Respiratory Health [MC_PC_19004] in partnership with SAIL Databank. BREATHE is funded through the UK Research and Innovation Industrial Strategy Challenge Fund and delivered through Health Data Research UK. The original study was supported by Health Research Council of New Zealand and Cure Kids, using Electronic monitoring devices provided by Nexus 6 (Auckland, New Zealand).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-72060-0>.

Correspondence and requests for materials should be addressed to H.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020