Pulmonology 29 (2023) 292—305

PULMONOLOGY

www.journalpulmonology.org

ORIGINAL ARTICLE

Identification by cluster analysis of patients with asthma and nasal symptoms using the MASK-air® mHealth app

J. Bousquet[a,b,c,*], B. Sousa-Pinto[d,e,f], J.M. Anto[g,h,i,j], R. Amaral[d,e,f], L. Brussino[k], G.W. Canonica[l,m], A.A. Cruz[n], B. Gemicioglu[o], T. Haahtela[p], M. Kupczyk[q], V. Kvedariene[r], D.E. Larenas-Linnemann[s], R. Louis[t], N. Pham-Thi[u], F. Puggioni[l,m], F.S. Regateiro[v], J. Romantowski[w], J. Sastre[x], N. Scichilone[y], L. Taborda-Barata[z], M.T. Ventura[aa], I. Agache[bb], A. Bedbrook[cc], K.C. Bergmann[a,b], S. Bosnic-Anticevich[dd], M. Bonini[ee,ff,gg], L.-P. Boulet[hh], G. Brusselle[ii], R. Buhl[jj], L. Cecchi[kk], D. Charpin[ll], C. Chaves-Loureiro[mm], W. Czarlewski[nn], F. de Blay[oo], P. Devillier[pp], G. Joos[ii], M. Jutel[qq], L. Klimek[rr], P. Kuna[q], D. Laune[ss], J.L. Pech[tt], M. Makela[p], M. Morais-Almeida[uu], R. Nadif[vv], M. Niedoszytko[w], K. Ohta[ww], N.G. Papadopoulos[xx], A. Papi[yy], D.R. Yeverino[zz], N. Roche[AA], A. Sá-Sousa[d,e,f], B. Samolinski[BB], M.H. Shamji[CC], A. Sheikh[DD], C. Suppli Ulrik[EE], O.S. Usmani[FF], A. Valiulis[GG], O. Vandenplas[HH], A. Yorgancioglu[II], T. Zuberbier[a,b], J.A. Fonseca[d,e,f]

[a] *Institute of Allergology, Charité — Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany*
[b] *Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Allergology and Immunology, Berlin, Germany*
[c] *University Hospital Montpellier, Montpellier, France*
[d] *MEDCIDS - Department of Community Medicine, Information and Health Decision Sciences; Faculty of Medicine, University of Porto, Porto, Portugal*
[e] *CINTESIS — Center for Health Technology and Services Research; University of Porto, Porto, Portugal*
[f] *RISE — Health Research Network; University of Porto, Porto, Portugal*
[g] *ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain*
[h] *IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain*
[i] *Universitat Pompeu Fabra (UPF), Barcelona, Spain*
[j] *CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain*
[k] *Department of Medical Sciences, Allergy and Clinical Immunology Unit, University of Torino & Mauriziano Hospital, Torino, Italy*
[l] *Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy*
[m] *Personalized Medicine, Asthma and Allergy, Humanitas Clinical and Research Center IRCCS, Rozzano, Italy*
[n] *Fundaçao ProAR, Federal University of Bahia and GARD/WHO Planning Group, Salvador, Bahia, Brazil*
[o] *Department of Pulmonary Diseases, Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey*
[p] *Skin and Allergy Hospital, Helsinki University Hospital, University of Helsinki, Finland*
[q] *Division of Internal Medicine, Asthma and Allergy, Barlicki University Hospital, Medical University of Lodz, Poland*
[r] *Institute of Clinical medicine, Clinic of Chest diseases and Allergology, Faculty of Medicine, Vilnius University and Institute of Biomedical Sciences, Department of Pathology, Faculty of Medicine, Vilnius University, Vilnius, Lithuania*
[s] *Center of Excellence in Asthma and Allergy, Médica Sur Clinical Foundation and Hospital, México City, Mexico*

* Corresponding author at: Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany.
E-mail address: jean.bousquet@orange.fr (J. Bousquet).

https://doi.org/10.1016/j.pulmoe.2022.10.005

[t] *Department of Pulmonary Medicine, CHU Liège and GIGA I3 research group, University of Liege, Belgium*
[u] *Ecole Polytechnique Palaiseau, IRBA (Institut de Recherche bio-Médicale des Armées), Bretigny, France*
[v] *Allergy and Clinical Immunology Unit, Centro Hospitalar e Universitário de Coimbra, and Coimbra Institute for Clinical and Biomedical Research (ICBR), Faculty of Medicine, University of Coimbra, and Institute of Immunology, Faculty of Medicine, University of Coimbra, Coimbra, Portugal*
[w] *Department of Allergology, Medical University of Gdańsk, Gdansk, Poland*
[x] *Fundacion Jimenez Diaz, CIBERES, Faculty of Medicine, Autonoma University of Madrid, Spain*
[y] *PROMISE Department, University of Palermo, Palermo, Italy*
[z] *Department of Immunoallergology, Cova da Beira University Hospital Centre, Covilhã, and UBIAir - Clinical & Experimental Lung Centre, and CICS-UBI Health Sciences Research Centre, University of Beira Interior, Covilhã, Portugal*
[aa] *Unit of Geriatric Immunoallergology, University of Bari Medical School, Bari, Italy*
[bb] *Transylvania University Brasov, Brasov, Romania*
[cc] *ARIA, Montpellier, France*
[dd] *Quality Use of Respiratory Medicine Group, Woolcock Institute of Medical Research, The University of Sydney, and Sydney Local Health District, Sydney, NSW, Australia*
[ee] *Department of Cardiovascular and Thoracic Sciences, Università Cattolica del Sacro Cuore, Rome, Italy*
[ff] *Department of Clinical and Surgical Sciences, Fondazione Policlinico Universitario A Gemelli IRCCS, Rome, Italy*
[gg] *National Heart and Lung Institute, Royal Brompton Hospital & Imperial College London, UK*
[hh] *Quebec Heart and Lung Institute, Laval University, Québec City, Quebec, Canada*
[ii] *Dept of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium*
[jj] *Dept of Pulmonary Medicine, Mainz University Hospital, Mainz, Germany*
[kk] *SOS Allergology and Clinical Immunology, USL Toscana Centro, Prato, Italy*
[ll] *Clinique des bronches, allergie et sommeil, Hôpital Nord, Marseille, France*
[mm] *Pneumology Unit, Hospitais da Universidade de Coimbra, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal*
[nn] *Medical Consulting Czarlewski, Levallois, France*
[oo] *Allergy Division, Chest Disease Department, University Hospital of Strasbourg, and Federation of translational medicine, University of Strasbourg, Strasbourg, France*
[pp] *VIM Suresnes, UMR 0892, Pôle des Maladies des Voies Respiratoires, Hôpital Foch, Université Paris-Saclay, Suresnes, France*
[qq] *Department of Clinical Immunology, Wrocław Medical University, and ALL-MED Medical Research Institute, Wroclaw, Poland*
[rr] *Department of Otolaryngology, Head and Neck Surgery, Universitätsmedizin Mainz, Mainz, and Center for Rhinology and Allergology, Wiesbaden, Germany*
[ss] *KYomed INNOV, Montpellier, France*
[tt] *University of Guadalajara, Guadalajara, Mexico*
[uu] *Allergy Center, CUF Descobertas Hospital, Lisbon, Portugal*
[vv] *Université Paris-Saclay, UVSQ, Univ. Paris-Sud, and Inserm, Equipe d'Epidémiologie Respiratoire Intégrative, CESP, Villejuif, France*
[ww] *National Hospital Organization, Tokyo National Hospital, and JATA Fukujuji Hospital, Tokyo, Japan*
[xx] *Allergy Department, 2nd Pediatric Clinic, University of Athens, Athens, Greece*
[yy] *Respiratory Medicine, Department of Translational Medicine, University of Ferrara, Ferrara, Italy*
[zz] *Allergy and Clinical Immunology Department, Hospital Universitario de Puebla, Puebla, Mexico*
[AA] *Pneumologie, AP-HP Centre Université de Paris Cité, Hôpital Cochin, Paris and UMR 1016, Institut Cochin, Paris, France*
[BB] *Department of Prevention of Environmental Hazards, Allergology and Immunology, Medical University of Warsaw, Poland*
[CC] *National Heart and Lung Institute, Imperial College London, and NIHR Imperial Biomedical Research Centre, London, UK*
[DD] *Usher Institute, The University of Edinburgh, Edinburgh, UK*
[EE] *Department of Respiratory Medicine, Copenhagen University Hospital-Hvidovre, and Institute of Cinical Medicine, University of Copenhagen, Denmark*
[FF] *National Heart and Lung Institute (NHLI), Imperial College London & Royal Brompton Hospital, Airways Disease Section, London, UK*
[GG] *Institute of Clinical Medicine and Institute of Health Sciences,Vilnius, and Medical Faculty of Vilnius University, Vilnius, Lithuania*
[HH] *Department of Chest Medicine, Centre Hospitalier Universitaire UCL, Namur, and Université Catholique de Louvain, Yvoir, Belgium*
[II] *Department of Pulmonary Diseases, Celal Bayar University, Faculty of Medicine, Manisa, Turkey*

Received 1 August 2022; accepted 13 October 2022
Available online 22 November 2022

KEYWORDS
Asthma;
Rhinitis;
Cluster analysis;
Treatment;
Control

Abstract
Background: The self-reporting of asthma frequently leads to patient misidentification in epidemiological studies. Strategies combining the triangulation of data sources may help to improve the identification of people with asthma. We aimed to combine information from the self-reporting of asthma, medication use and symptoms to identify asthma patterns in the users of an mHealth app.
Methods: We studied MASK-air® users who reported their daily asthma symptoms (assessed by a 0-100 visual analogue scale – "VAS Asthma") at least three times (either in three different months or in any period). K-means cluster analysis methods were applied to identify asthma patterns based on: (i) whether the user self-reported asthma; (ii) whether the user reported asthma medication use and (iii) VAS asthma. Clusters were compared by the number of medications used, VAS asthma levels and Control of Asthma and Allergic Rhinitis Test (CARAT) levels.
Findings: We assessed a total of 8,075 MASK-air® users. The main clustering approach resulted in the identification of seven groups. These groups were interpreted as probable: (i) severe/uncontrolled asthma despite treatment (11.9-16.1% of MASK-air® users); (ii) treated and partly-controlled asthma (6.3-9.7%); (iii) treated and controlled asthma (4.6-5.5%); (iv) untreated uncontrolled asthma (18.2-20.5%); (v) untreated partly-controlled asthma (10.1-10.7%); (vi) untreated controlled asthma (6.7-8.5%) and (vii) no evidence of asthma (33.0-40.2%). This classification was validated in a study of 192 patients enrolled by physicians.
Interpretation: We identified seven profiles based on the probability of having asthma and on its level of control. mHealth tools are hypothesis-generating and complement classical epidemiological approaches in identifying patients with asthma.

Introduction

Self-reporting is a common method for gathering data in medical research. While self-reported data may be prone to information bias,[1] they can help to complement other data collection approaches.[2] Relying on the self-reporting of asthma may be problematic, as patients self-report bronchoconstriction variably,[3] may not have been diagnosed (asthma under-diagnosis ranges between 19-73%[4,5]) or believe they do not have asthma despite being symptomatic.[6,7]

Cluster analysis, combining information from different variables, may help to overcome undue reliance on self-reported asthma, improving the identification and characterisation of patients with asthma. This approach has been used to understand the heterogeneity of asthma[8-10] or to test different hypotheses in adult patients with asthma.[8,11,12] The application of clustering approaches to asthma real-world data (RWD) may also be valuable. As an example, RWD obtained with MASK-air® (Mobile Airways Sentinel networK), a validated mobile app for rhinitis and asthma, have enabled the definition of new phenotypes of allergic rhinitis[13] and the assessment of adherence to treatment.[14] MASK-air® may result in similar advances in asthma, but the correct identification of asthmatic patients is required.

In this study, we used cluster analysis to identify and characterise asthma patterns amongst MASK-air® users in a non-supervised way. We aimed to understand whether RWD from mobile apps can be informative for the identification of asthma, hinting at the frequency of misdiagnosis and, potentially, mistreatment.

Methods

Study design

We performed a cross-sectional analysis using the MASK-air® database to identify asthma patterns, assessing three different samples (Supplementary Figure 1). We performed cluster analysis to identify asthma patterns based on the self-reporting of asthma, asthma medication use and VAS asthma, adopting a stepwise approach to check for consistency of results. We compared the characteristics of the obtained clusters and we validated them in a sample of patients in whom asthma diagnosis had been assessed by a physician during a transfer of innovation (Twinning) of the European Innovation Partnership on Active and Healthy Ageing.[15]

Setting and participants

MASK-air®, available since 2015, can be downloaded via the Apple App and Google Play Stores. We assessed three samples of MASK-air® users from May 2015 to December 2020. The users were aged 16-90 years and had self-reported allergic rhinitis. Samples 1 and 2 consisted of all MASK-air® users reporting VAS asthma in at least three different months - to limit the possibility of having "false-positives" (e.g., patients with high values of VAS asthma or those using asthma medication inappropriately within short periods of time as a result of respiratory infections or other non-asthma-related causes). In Sample 1, only users who answered to the Control of Allergic Rhinitis and Asthma Test (CARAT)[16] at least once were included. In Sample 2, all users

were included irrespective of having answered to CARAT or not. Sample 3 consisted of all MASK-air® users reporting at least three VAS asthma, irrespective of the timing.

In the Twinning project, patients were enrolled during a medical consultation with an asthma specialist (14 centres from Germany, Italy, Lithuania, Poland, Portugal and Spain) and were instructed to use MASK-air®.[15] Asthma was diagnosed according to the Global Initiative for Asthma (GINA),[17] with patients having a pulmonary function test. Participants were classified as having "current asthma", "past asthma" or "no current or past asthma".

Ethics

MASK-air® follows the GDPR regulations.[18] All data are anonymised using k-anonymity. An independent Review Board (Bohn-Köln; 11.05.2017; N°17-069) approval was obtained for the MASK-air studies.[15] For the Twinning project, additional local review board approvals were obtained (Mannheim − reference: 2018-527N-MA, 29.03.2018 for Germany; Coimbra − reference: CHUC-022-18, 14.09.2018 for Portugal; Warsaw − reference: AKBE/213/2019, 13.05.2019 for Poland; Vilnius 2021 for Lithuania; Bari − reference: 7287, 30.03.2022 for Italy). For patients who did not participate in the Twinning, individual boards in different countries were not required since users agree to the analysis of their data in the terms of use.

Data sources and variables

MASK-air® comprises a daily monitoring questionnaire assessing (i) the daily impact of asthma and rhinitis symptoms by means of 0-100 VASs and (ii) users' asthma and rhinitis daily medications (available from country-specific lists with prescribed and over-the-counter medications).

MASK-air® also allows users to answer to CARAT, a 10-item questionnaire assessing rhinitis and asthma control in the previous four weeks.[19] We considered "CARAT asthma" to correspond to questions 5-7 ("Shortness of breath/dyspnoea", "Wheezing in the chest" and "Chest tightness upon physical exercise"), with a score of $\leq$6 out of 9 indicating symptoms suggestive of asthma.

Size of the study

Data from all users meeting the inclusion criteria were analysed.

Biases

There are potential information biases related to the self-reported nature of the data collection. There may be an over-representation of users suffering from moderate-to-severe asthma[20] and of younger individuals. Additionally, it is not known whether users fill in the MASK-air® daily questionnaire before or after treatment for a given day.

Data analysis

A full description of the data analysis methods is available in the Supplement. In brief, in each sample, we applied k-means cluster analysis methods to identify patterns of MASK-air® users according to self-reported asthma, use of asthma medication and VAS asthma (supplementary Figure 2). Obtained clusters were assessed and compared regarding asthma- and rhinitis-related variables as well as patients' demographic characteristics. To check for consistency of results, we compared clusters obtained by the main clustering approach with those obtained using alternative approaches, and in a sample of patients with physician-diagnosed asthma (Twinning participants).

Results

Demographic and clinical characteristics

Among the 17,780 patients of the MASK-air® database, 8,075 provided data on VAS asthma at least three different times (Sample 3). Of those, 3,797 provided VAS asthma in at least three different months (Sample 2), including 466 patients who answered to CARAT at least once (Sample 1) (Supplementary Figure 3). The demographic characteristics of patients are available in Supplementary Table 1.

Cluster analysis results

Main analysis approach

An optimal number of four clusters (A-D) was identified in the patients of Sample 1 (Table 1A):

- Cluster A: 96% of the patients self-reported asthma and 91% reported $\geq$3 days of asthma medication. VAS asthma values were high (median maximum value=85/100). Asthma symptoms identified by "CARAT-asthma" were observed in 67% of the patients.
- Cluster B: 93% of the patients self-reported asthma and 87% reported $\geq$3 days of asthma medication. Maximum VAS asthma values were moderate (median=45). Asthma symptoms identified by "CARAT-asthma" were observed in 32% of the patients.
- Cluster C: 50% of the patients self-reported asthma and most never reported any asthma medication. High maximum VAS asthma values were reported (median=74). Asthma symptoms identified by "CARAT-asthma" were observed in 58% of the patients.
- Cluster D: Few patients self-reported asthma (15%), most never reported any asthma medication (97%) and VAS maximum asthma values were low (median=11). Asthma symptoms identified by "CARAT-asthma" were observed in 15% of the patients.

The same optimal number of clusters was identified in Samples 2 and 3. The characteristics of the four clusters were highly consistent across all samples (Tables 1B and 1C).

We subsequently identified two subgroups within Cluster B and three subgroups within Cluster D. The two subgroups of Cluster B differed on VAS asthma (Table 2; Supplementary Table 2). The three subgroups of Cluster D included (i) one subgroup with a low frequency of asthma self-reporting ($<$20%) and moderate maximum VAS asthma values; (ii) one subgroup with all participants self-reporting asthma and

Table 1 Description of the four asthma-related clusters using the k-means approach.

A. Sample 1: Patients with at least 3 VAS asthma in 3 different months who answered at least once to CARAT

	Cluster A	Cluster B	Cluster C	Cluster D	*p*-value
N (%)	75 (16.1)	69 (14.8)	90 (19.3)	232 (49.8)	
Reported days — *N* (average days per user)	8888 (118.5)	9066 (131.4)	7646 (85.0)	21,730 (93.7)	
Females*	62 (82.7)	46 (66.7)	58 (64.4)	147 (63.4)	0.019[a]
Age‖	41.1 (11.2)	40.7 (11.4)	39.2 (14.0)	37.5 (13.6)	0.104
Self-reported asthma*	72 (96.0)	64 (92.8)	45 (50.0)	35 (15.1)	<0.001
Asthma medication reporting*					<0.001
0 days	0	0	79 (87.8)	226 (97.4)	
1 day	0	0	11 (12.2)	6 (2.6)	
2 days	7 (9.3)	9 (13.0)	0	0	
3 or more days	68 (90.7)	60 (87.0)	0	0	
Total days reporting asthma medication*					
SABA	1379 (15.5)	578 (6.4)	9 (0.1)	7 (0.03)	<0.001
LABA+ICS	3916 (44.1)	3369 (37.2)	8 (0.1)	2 (0.01)	<0.001
ICS	1168 (13.1)	1443 (15.9)	3 (0.04)	0	<0.001
OCS[b]	507 (5.7)	41 (0.5)	61 (0.8)	31 (0.1)	
LAMA	651 (7.3)	456 (5.1)	0	0	
Omalizumab	7 (0.1)	6 (0.1)	0	0	
VAS asthma					
Maximum value†	85 (76-94)	45 (30-55)	74 (61-86)	11 (3-26)	<0.001
Three highest values†	73 (64-83)	35 (23-45)	61 (48-75)	6 (1-14)	<0.001
Days with VAS asthma >50*	1392 (15.7)	35 (0.4)	1057 (13.8)	17 (0.1)	<0.001
Maximum VAS dyspnea†	68 (56-83)	20 (4-41)	59 (34-74)	20 (7-36)	<0.001
CARAT asthma (questions 5-7)†	5 (2–7)	7 (6–9)	6 (4–8)	9 (7–9)	<0.001
Presence of asthma symptoms[c] *	50 (66.7)	22 (31.9)	52 (57.8)	36 (15.5)	<0.001
CARAT (questions 1–10)†	13 (8–16)	19 (17–23)	15 (11–19)	20 (16–24)	<0.001
Uncontrolled[d] *	73 (97.3)	53 (76.8)	81 (90.0)	174 (75.0)	<0.001
Maximum CSMS†	68 (59–78)	36 (30–46)	63 (46–69)	39 (20–54)	<0.001
Maximum VAS global†	84 (74–96)	49 (41–65)	87 (71–100)	65 (44–84)	<0.001
Maximum VAS eyes†	78 (60–92)	40 (19–59)	76 (64–97)	50 (27–76)	<0.001
Maximum VAS nose†	86 (70–98)	58 (42–75)	88 (75–100)	69 (44–89)	<0.001
Maximum VAS work†	61 (43–73)	27 (10–46)	62 (44–83)	31 (10–54)	<0.001
Maximum VAS sleep†	87 (71–98)	67 (44–84)	86 (70–100)	66 (41–86)	<0.001
Total days reporting rhinitis medication*					
Oral antihistamines monotherapy	1199 (13.5)	934 (10.3)	710 (9.3)	2440 (11.2)	<0.001
Intranasal steroids monotherapy	361 (4.1)	768 (8.5)	378 (4.9)	776 (3.6)	<0.001
Azelastine–fluticasone monotherapy	346 (3.9)	543 (6.0)	107 (1.4)	1009 (4.6)	<0.001
Oral antihistamines + intranasal steroids	2087 (23.5)	1721 (19.0)	454 (5.9)	1009 (4.6)	<0.001
Azelastine–fluticasone + other rhinitis medication	878 (9.9)	850 (9.4)	125 (1.6)	520 (2.4)	<0.001
Conjunctivitis*	68 (90.7)	49 (71.0)	66 (73.3)	183 (78.9)	0.016[a]
Sensitisation[e] *					0.181
Monosensitisation[e]	8 (11.4)	6 (8.8)	8 (9.1)	31 (13.7)	
Polysensitisation[e]	51 (72.9)	40 (58.8)	51 (58.0)	132 (58.1)	

B. Sample 2: All patients with at least 3 VAS asthma in 3 different months

	Cluster A	Cluster B	Cluster C	Cluster D	*p*–value
N (%)	451 (11.9)	414 (10.9)	780 (20.5)	2152 (56.7)	
Reported days — *N* (average days per user)	38,823 (86.1)	35,723 (86.3)	47,352 (60.7)	134,941 (62.7)	
Females*	310 (68.7)	234 (56.5)	460 (59.0)	1138 (52.9)	<0.001
Age‖	41.1 (14.3)	40.1 (14.1)	38.3 (13.8)	35.5 (13.2)	<0.001
Self–reported asthma*	432 (95.8)	389 (94.0)	391 (50.1)	341 (15.8)	<0.001
Asthma medication reporting*					<0.001
0 days	0	0	698 (89.5)	2102 (97.7)	
1 day	4 (0.9)	10 (2.4)	82 (10.5)	50 (2.3)	
2 days	68 (15.1)	64 (15.5)	0	0	
3 or more days	379 (84.0)	340 (82.1)	0	0	
Total days reporting asthma medication*					
SABA	4285 (11.0)	1586 (4.4)	66 (0.1)	37 (0.03)	<0.001
LABA+ICS	16,275 (41.9)	15,038 (42.1)	74 (0.2)	23 (0.02)	<0.001
ICS	4658 (12.0)	5722 (16.0)	25 (0.1)	22 (0.02)	<0.001
OCS[b]	1331 (3.4)	243 (0.7)	244 (0.5)	141 (0.1)	
LAMA	1453 (3.7)	534 (1.5)	0	0	
Biologics	112 (0.3)	86 (0.2)	0	0	
VAS asthma					
Maximum value†	81 (69–92)	38 (25–49)	72 (58–85)	8 (2–22)	<0.001
Three highest values†	69 (58–82)	27 (15–37)	56 (44–72)	4 (0–12)	<0.001
Days with VAS asthma >50*	5610 (14.5)	91 (0.3)	4799 (10.1)	94 (0.1)	<0.001
Maximum VAS dyspnea†	69 (54–82)	31 (18–45)	61 (42–75)	19 (7–34)	<0.001
Maximum CSMS†	63 (52–72)	36 (25–47)	62 (50–71)	37 (26–53)	<0.001
Maximum VAS global†	80 (69–93)	49 (34–67)	81 (68–95)	61 (39–81)	<0.001
Maximum VAS eyes†	71 (51–89)	34 (20–60)	75 (57–90)	44 (21–71)	<0.001
Maximum VAS nose†	82 (67–95)	53 (34–75)	85 (70–100)	66 (41–85)	<0.001
Maximum VAS work†	57 (37–71)	26 (9–43)	58 (40–74)	29 (10–52)	<0.001
Maximum VAS sleep†	72 (26–90)	52 (33–77)	79 (60–94)	56 (34–79)	<0.001
Total days reporting rhinitis medication*					
Oral antihistamines monotherapy	4594 (11.8)	3852 (10.8)	4984 (10.5)	16,971 (12.6)	<0.001
Intranasal steroids monotherapy	1787 (4.6)	3864 (10.8)	2290 (4.8)	7290 (5.4)	<0.001
Azelastine–fluticasone monotherapy	1465 (3.8)	1217 (3.4)	1288 (2.7)	5270 (3.9)	<0.001
Oral antihistamines + intranasal steroids	5949 (15.3)	3362 (9.4)	2982 (6.3)	8158 (6.0)	<0.001

Azelastine−fluticasone + other rhinitis medication	2568 (6.6)	1804 (5.0)	1601 (3.4)	3244 (2.4)	<0.001
Conjunctivitis*	341 (75.6)	293 (70.8)	590 (75.6)	1581 (73.5)	0.239
Sensitisation[f] *					0.149
Monosensitisation[f]	18 (6.3)	20 (7.4)	36 (7.8)	97 (7.6)	
Polysensitisation[f]	136 (47.7)	113 (41.7)	181 (39.3)	486 (38.1)	

C. Sample 3: All patients with at least 3 VAS asthma

	Cluster A	Cluster B	Cluster C	Cluster D	*p*−value
N (%)	957 (11.9)	937 (11.6)	1468 (18.2)	4713 (58.4)	
Reported days − *N* (average days per user)	52,649 (55.0)	44,468 (47.5)	54,438 (37.1)	145,614 (30.9)	
Females*	675 (70.5)	562 (60.0)	907 (61.8)	2554 (54.2)	<0.001
Age‖	39.5 (13.5)	38.3 (13.8)	37.1 (13.6)	34.6 (12.9)	<0.001
Self−reported asthma*	875 (91.4)	754 (80.5)	680 (46.3)	763 (16.2)	<0.001
Asthma medication reporting*					<0.001
0 days	0	0	1316 (89.6)	4604 (97.7)	
1 day	6 (0.6)	0	152 (10.4)	109 (2.3)	
2 days	82 (8.6)	117 (12.5)	0	0	
3 or more days	869 (90.8)	820 (87.5)	0	0	
Total days reporting asthma medication*					
SABA	5531 (10.5)	1581 (3.6)	65 (0.1)	39 (0.03)	<0.001
LABA+ICS	20,320 (38.6)	14,135 (31.8)	54 (0.1)	28 (0.02)	<0.001
ICS	5471 (10.4)	5853 (13.2)	25 (0.1)	13 (0.01)	<0.001
OCS[b]	1480 (2.8)	264 (0.6)	355 (0.7)	231 (0.2)	
LAMA	1901 (3.6)	307 (0.7)	1 (0.002)	0	
Biologics	116 (0.2)	88 (0.2)	1 (0.002)	0	
VAS asthma					
Maximum value†	78 (65−92)	30 (13−45)	69 (54−84)	6 (0−18)	<0.001
Three highest values†	65 (53−79)	18 (6−30)	52 (39−68)	2 (0−9)	<0.001
Days with VAS asthma >50*	7677 (14.6)	154 (0.3)	6001 (11.0)	154 (0.1)	<0.001
Maximum VAS dyspnea†	67 (53−83)	29 (16−41)	61 (46−76)	17 (7−32)	<0.001
CARAT asthma (questions 5−7)[g] †	6 (3−8)	7 (6−9)	6 (5−8)	9 (7−9)	<0.001
Presence of asthma symptoms [c,g] *	125 (74.9)	55 (44.7)	90 (61.6)	78 (19.6)	<0.001
CARAT (questions 1−10)[g] †	15 (10−18)	19 (15−24)	15 (12−20)	20 (15−23)	<0.001
Uncontrolled[d,g] *	159 (95.2)	107 (87.0)	138 (94.5)	333 (83.7)	<0.001
Maximum CSMS†	61 (50−70)	34 (23−47)	60 (49−71)	35 (25−51)	<0.001
Maximum VAS global†	78 (66−92)	49 (32−69)	78 (66−92)	58 (36−78)	<0.001
Maximum VAS eyes†	66 (43−84)	32 (13−59)	72 (51−88)	40 (16−68)	<0.001
Maximum VAS nose†	78 (64−93)	52 (32−75)	81 (66−97)	62 (37−82)	<0.001
Maximum VAS work†	52 (24−67)	21 (4−40)	53 (27−70)	25 (4−49)	<0.001
Maximum VAS sleep†	79 (63−95)	54 (32−78)	77 (58−92)	55 (33−77)	<0.001
Total days reporting rhinitis medication*					
Oral antihistamines monotherapy	5880 (11.2)	4243 (9.5)	6431 (11.8)	19,395 (13.3)	<0.001
Intranasal steroids monotherapy	3060 (5.8)	4039 (9.1)	2976 (5.5)	8405 (5.8)	<0.001
Azelastine−fluticasone monotherapy	1963 (3.7)	1028 (2.3)	1476 (2.7)	6145 (4.2)	<0.001
Oral antihistamines + intranasal steroids	7600 (14.4)	4693 (10.6)	3542 (6.5)	7967 (5.5)	<0.001
Azelastine−fluticasone + other rhinitis medication	3773 (7.2)	1766 (4.0)	1487 (2.7)	3414 (2.3)	<0.001
Conjunctivitis*	717 (74.9)	660 (70.4)	1136 (77.4)	3487 (74.0)	0.002[a]
Sensitisation[h] *					0.021[a]
Monosensitisation[h]	33 (10.5)	28 (10.1)	38 (10.4)	121 (12.0)	
Polysensitisation[h]	195 (62.3)	185 (66.5)	209 (57.4)	657 (65.4)	

CARAT: Control of Allergic Rhinitis and Asthma Test; CSMS: Combined symptom-medication score; ICS: Inhaled corticosteroid; IQR: Interquartile range; LABA: Long-acting beta-agonist; LAMA: Long-acting muscarinic antagonist; OCS: Oral corticosteroid; SABA: Short-acting beta-agonist; VAS: Visual Analogue Scale.

* Results presented as *N*(%).

‖ Results presented as mean (SD).

† Results presented as median (percentile 25-percentile 75).

[a] Non-significant after applying the Bonferroni correction.

[b] It is not possible to differentiate OCS used for asthma or for allergic rhinitis.

[c] Score ≤6.

[d] Score ≤24.

[e] Number of patients for whom sensitisation data are available: 70 for cluster A, 68 for cluster B, 88 for cluster C, and 227 for cluster D.

[f] Number of patients for whom sensitisation data are available: 285 for cluster A, 271 for cluster B, 460 for cluster C, and 1275 for cluster D.

[g] Number of days reporting CARAT: 167 for cluster A, 123 for cluster B, 146 for cluster C, and 398 for cluster D.

[h] Number of patients for whom sensitisation data are available: 313 for cluster A, 278 for cluster B, 364 for cluster C, and 1005 for cluster D.

with low maximum VAS asthma values and (iii) one subgroup with no participants self-reporting asthma and with very low VAS asthma values. For Clusters A and C, the silhouette score was <0.5, suggesting that clustering may not be adequate. Nevertheless, since there were around 50% of patients self-reporting asthma in Cluster C, we performed an ancillary analysis comparing Cluster C patients with self-reported asthma (C') versus those with no self-reported asthma (C''). Overall, patients of the two subgroups were similar (Supplementary Table 3).

Since selecting patients reporting VAS asthma in at least three different months could be interpreted as having some degree of arbitrariness, we performed sensitivity analyses applying the same methods in patients reporting VAS asthma

Table 2 Asthma-related clusters and respective subgroups obtained using a two-step k-means (Sample 2).

	Cluster A ("Treated uncontrolled asthma")	Cluster B1 ("Treated partly-controlled asthma")	Cluster B2 ("Treated controlled asthma")	Cluster C ("Untreated uncontrolled asthma")	Cluster D1 ("Untreated partly- controlled asthma")	Cluster D2 ("Untreated controlled asthma")	Cluster D3 ("No evidence of asthma")
N (%)	451 (11.9)	239 (6.3)	175 (4.6)	780 (20.5)	406 (10.7)	323 (8.5)	1423 (37.5)
Reported days — *N*	38,823	23,953	11,770	47,352	30,907	16,287	87,747
Average days per user - *N*	86.1	100.2	67.3	60.7	76.1	50.4	61.7
Females*	310 (68.7)	138 (57.7)	96 (54.9)	460 (59.0)	209 (51.5)	176 (54.5)	753 (52.9)
Age‖	41.1 (14.3)	40.8 (14.5)	39.2 (13.6)	38.3 (13.8)	37.1 (13.0)	36.3 (13.9)	34.8 (13.1)
Self-reported asthma*	432 (95.8)	228 (95.4)	161 (92.0)	391 (50.1)	18 (4.4)	323 (100)	0
Asthma medication reporting*							
0 days	0	0	0	698 (89.5)	401 (98.8)	284 (87.9)	1417 (99.6)
1 day	4 (0.9)	10 (4.2)	0	82 (10.5)	5 (1.2)	39 (12.1)	6 (0.4)
2 days	68 (15.1)	31 (13.0)	33 (18.9)	0	0	0	0
3 days or more	379 (84.0)	198 (82.8)	142 (81.1)	0	0	0	0
Total days reporting asthma medication*							
SABA	4285 (11.0)	1180 (4.9)	406 (3.4)	66 (0.1)	4 (0.01)	29 (0.2)	4 (0.01)
LABA+ICS	16,275 (41.9)	9508 (39.7)	5530 (47.0)	74 (0.2)	0	22 (0.1)	1 (0.001)
ICS	4658 (12.0)	3194 (13.3)	2528 (21.5)	25 (0.1)	4 (0.01)	16 (0.1)	2 (0.002)
OCS [a]	1331 (3.4)	206 (0.9)	37 (0.3)	244 (0.5)	9 (0.03)	8 (0.1)	124 (0.1)
LAMA	1453 (3.7)	465 (1.9)	69 (0.6)	0	0	0	0
Biologics	112 (0.3)	81 (0.3)	5 (0.04)	0	0	0	0
VAS asthma							
Maximum value†	81 (69—92)	47 (41—54)	21 (12—29)	72 (58—85)	36 (26—49)	20 (7-31)	4 (1—9)
Three highest values†	69 (58—82)	35 (30—43)	13 (6-20)	56 (44—72)	19 (13—25)	12 (3-20)	1 (0—5)
Days with VAS asthma >50 *	5610 (14.5)	90 (0.4)	1 (0.01)	4799 (10.1)	89 (0.3)	5 (0.03)	0
Maximum VAS dyspnea †	69 (54—82)	38 (26—49)	17 (12—27)	61 (42—75)	29 (16—40)	20 (13—33)	10 (5-23)
CARAT asthma (questions 5-7)[b] †	5 (2—7)	8 (7—9)	7 (6—8)	6 (4—8)	9 (7—9)	9 (7—9)	9 (8—9)
Presence of asthma symptoms[b,c,d,*]	50 (66.7)	19 (42.2)	8 (33.3)	52 (57.8)	9 (19.1)	7 (22.6)	20 (13.0)
CARAT (questions1-10) [b,†]	13 (8-16)	20 (16—25)	21 (19—23)	15 (11—19)	20 (16—25)	20 (17—25)	20 (16—25)
Uncontrolled [b,d,*]	73 (97.3)	36 (80.0)	21 (87.5)	81 (90.0)	34 (72.3)	23 (74.2)	117 (76.0)
Maximum CSMS†	63 (52—72)	42 (20)	28 (20)	62 (50—71)	46 (35—60)	29 (22—46)	35 (24—51)
Maximum VAS global†	80 (69—93)	53 (42—71)	40 (21—60)	81 (68—95)	72 (51—86)	47 (29—68)	61 (38—81)
Maximum VAS eyes†	71 (51—89)	42 (28—66)	29 (12—50)	75 (57—90)	57 (38—78)	34 (14—55)	42 (19—70)
Maximum VAS nose†	82 (67—95)	59 (42—79)	46 (27—66)	85 (70-100)	76 (55—91)	51 (32—75)	66 (40—85)
Maximum VAS work†	57 (37—71)	31 (14—48)	16 (5-31)	58 (40—74)	42 (21—60)	21 (6-40)	28 (9-52)
Maximum VAS sleep†	72 (26.90)	61 (40—82)	45 (18—64)	79 (60—94)	53 (14—76)	50 (26—75)	56 (33—78)
Total days reporting rhinitis medication*							
Oral antihistamines monotherapy	4594 (11.8)	2864 (12.0)	988 (8.4)	4984 (10.5)	4780 (15.5)	1165 (7.2)	11,026 (12.6)
Intranasal steroids monotherapy	1787 (4.6)	2291 (9.6)	1573 (13.4)	2290 (4.8)	1999 (6.5)	681 (4.2)	4610 (5.3)
Azelastine-fluticasone monotherapy	1465 (3.8)	908 (3.8)	309 (2.6)	1288 (2.7)	1220 (3.9)	346 (2.1)	3704 (4.2)

Table 2 (*Continued*)

	Cluster A ("Treated uncontrolled asthma")	Cluster B1 ("Treated partly-controlled asthma")	Cluster B2 ("Treated controlled asthma")	Cluster C ("Untreated uncontrolled asthma")	Cluster D1 ("Untreated partly- controlled asthma")	Cluster D2 ("Untreated controlled asthma")	Cluster D3 ("No evidence of asthma")
Oral antihistamines + intranasal steroids	5949 (15.3)	2263 (9.4)	1099 (9.3)	2982 (6.3)	1637 (5.3)	1165 (7.2)	5356 (6.1)
Azelastine-fluticasone + other rhinitis medication	2568 (6.6)	1448 (6.0)	356 (3.0)	1601 (3.4)	1280 (4.1)	348 (2.1)	1616 (1.8)
Conjunctivitis [*]	341 (75.6)	171 (71.5)	122 (69.7)	590 (75.6)	300 (73.9)	235 (72.8)	1046 (73.5)
Sensitisation [e,*]							
Monosensitisation [e]	18 (6.3)	13 (8.3)	7 (6.1)	36 (7.8)	14 (5.5)	10 (5.1)	73 (8.9)
Polysensitisation [e]	136 (47.7)	65 (41.7)	48 (41.7)	181 (39.3)	101 (39.6)	72 (36.4)	313 (38.1)

CARAT: Control of Allergic Rhinitis and Asthma Test; CSMS: Combined symptom-medication score; ICS: Inhaled corticosteroid; IQR: Interquartile range; LABA: Long-acting beta-agonist; SABA: Short-acting beta-agonist; VAS: Visual Analogue Scale.

[*] Results presented as *N*(%).

[‖] Results presented as mean (SD).

[†] Results presented as median (percentile 25-percentile 75).

[a] It is not possible to differentiate OCS used for asthma or for allergic rhinitis.

[b] Number of patients reporting CARAT: 75 for cluster A, 45 for cluster B1, 24 for cluster B2, 90 for cluster C, 47 for cluster D1, 31 for cluster D2, and 154 for cluster D3.

[c] Score $\leq$6.

[d] Score $\leq$24.

[e] Number of patients for whom sensitisation data are available: 285 for cluster A, 156 for cluster B1, 115 for cluster B2, 460 for cluster C, 255 for cluster D1, 198 for cluster D2, and 822 for cluster D3.

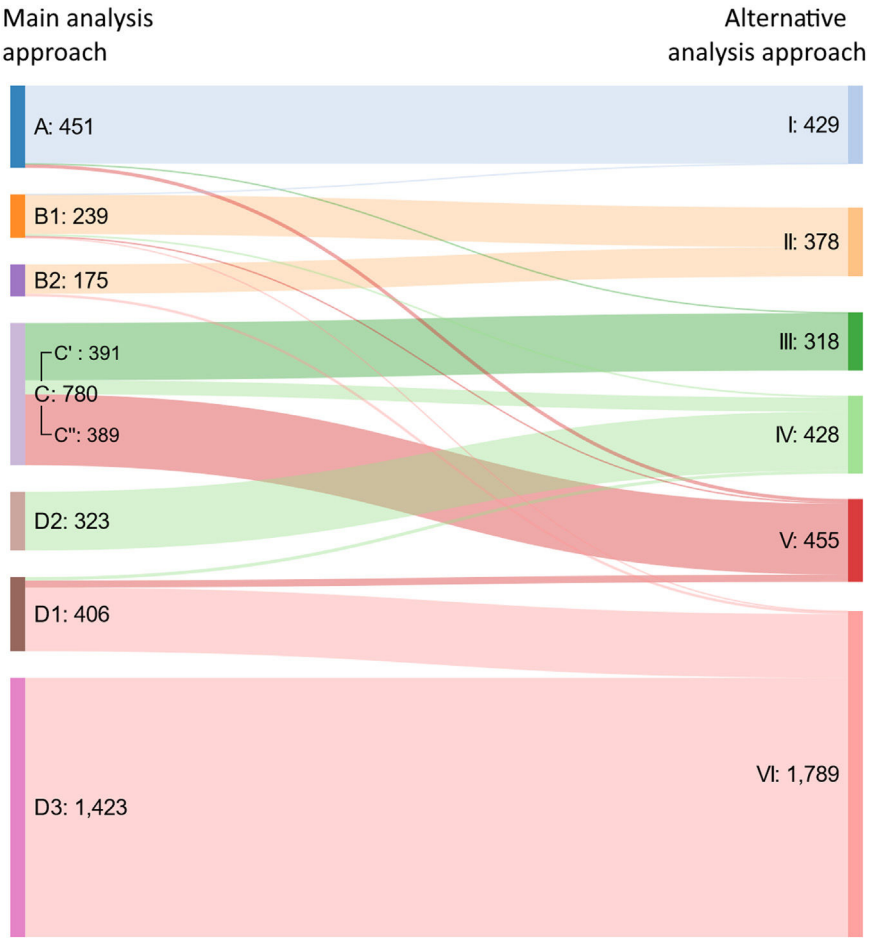

Fig. 1 Classification of patients who reported VAS asthma in at least three different months (Sample 2) with clustering based on the main analysis approach and on the alternative analysis approach of this study.

in at least four and five different months. Similar results were obtained.

Alternative analysis approach

Four clusters were identified among patients self-reporting asthma, while two clusters were identified among those not self-reporting asthma (Supplementary Tables 5-6). Using a Sankey diagram, the two approaches showed consistent results (Fig. 1).

Phenotypic characteristics of the clusters

Median VAS asthma maximal levels were over 50/100 for Clusters A and C, indicating "uncontrolled asthma". VAS asthma levels ranged from 20 to 49/100 in Clusters B1 and D1 (indicating "partly-controlled asthma") and were under 20/100 in Clusters B2 and D2 (indicating "controlled asthma"). The lowest levels were in Cluster D3 (Supplementary Figure 4).

Patients were mostly undertreated in Clusters C, D1, D2 and D3. In Cluster C, only half of the patients self-reported asthma. Therefore, Clusters C and D1 may include patients with under-diagnosed asthma. A possible clinical interpretation of the seven clusters observed with the main approach is available in Table 3.

Throughout the different months of the year, the order of VAS asthma levels was found to be consistent, with the highest levels being observed in Cluster A, followed by C, B1 and the remaining groups (Supplementary Figure 5).

Besides differences in asthma features, the seven clusters differed in the participants' demographics, in the VASs on allergy symptoms and in rhinitis treatment (Table 2, Supplementary Figures 4 and 6). The reported rhinitis treatments varied between clusters, ranging from 22.8-42.1% of days. Co-medication was reported in 21.9% of days for Cluster A, 15.4% for Cluster B1, 12.3% for Cluster B2 and around 9-10% of days in untreated asthma clusters.

Validation of the cluster classification

We analysed 192 Twinning participants, comparing the cluster classification obtained by the main analysis approach with physician-diagnosed asthma (Supplementary Table 7).

Patients clustered as having "probable asthma" (clusters A, D and C') had a physician diagnosis of current or past asthma in 92.3% of cases. Patients with "no evidence of current asthma" (cluster D3) had a diagnosis of "no current

Table 3 Clinical interpretation of the clusters obtained following clustering approaches.

Asthma			Main clustering approach		Alternative clustering approach		Clinical interpretation
Treatment	Control	Majority of self-report	Cluster	% of users[a]	Cluster	% of users[a]	
Treated	Uncontrolled	Yes	A	11.9-16.1	I	10.5-15.5	• Probable asthma: Treated uncontrolled asthma
	Partly- controlled		B1	6.3-9.7	II	9.6-13.7	• Probable asthma: Treated partly-controlled asthma
	Controlled		B2	4.6-5.5			• Probable asthma: Treated and controlled asthma
Untreated	Uncontrolled	Yes	C'[b]	9.7-10.2	III	7.0-8.6	• Probable asthma: Untreated and uncontrolled asthma with self-reported asthma (possible undertreated asthma)
		No	C'' [b]	8.4-10.3	V	11.7-12.9	• Possible asthma: Untreated uncontrolled asthma with no self-reported asthma (possible underdiagnosed asthma)
	Partly- controlled		D1	10.1-10.7	VI	*40.8-50.3* [c]	• Possible asthma: Untreated partly-controlled asthma (possible under-diagnosed asthma)
	Controlled	Yes	D2	6.7-8.5	IV	8.6-11.3	• Possible asthma: Untreated controlled asthma (possible over-diagnosed asthma or asthma in clinical remission)
		No	D3	33.0-40.2	VI	*40.8-50.3* [c]	• No evidence of current asthma

[a] Range of percentages across the three samples.
[b] Cluster C was divided by design (not by unsupervised learning approach).
[c] Range of percentages of cluster VI as a whole.

asthma" in 90.4% of cases. Patients with "uncontrolled underdiagnosed asthma" (cluster C") had an infrequent physician diagnosis of asthma, supporting the label of underdiagnosis.

A patient with current asthma displayed an 85.5% probability of being classified in a cluster of probable asthma (sensitivity) and a 93.4% probability of being in a cluster of probable or possible asthma. A patient with no history of asthma displayed a 52.6% probability of being classified as having no asthma (specificity) and 79.3% as having current asthma.

The classification of probable *versus* possible or no asthma for the identification of current asthma *versus* past or no asthma displays an agreement of 81% and a kappa coefficient of 0.610.

Discussion

Cluster analysis approaches were used to identify asthma control patterns in MASK-air® users combining information from self-reported asthma status, reported asthma medication use and VAS asthma. We identified seven profiles of asthma control and treatment patterns. These profiles were replicated in three samples and were validated in a sub-sample of physician-assessed patients.

Limitations and strengths

This study has some limitations. First, clustering was not performed based on patients from asthma clinics with a confirmed diagnosis of asthma. This type of study (i) would have a limited number of patients and (ii) would have mostly included severe patients and patients under treatment. However, we validated the results of the cluster classification in a sample of participants with a physician-diagnosis of asthma. Further information biases may occur, resulting from incorrect information on self-reported asthma or medication use. However, the consistency of the results suggests that this is unlikely.

All assessed patients displayed self-reported rhinitis, and the results are only valid for those with nasal symptoms. These patients do however represent a very large proportion of patients with asthma. Furthermore, there may be an over-representation of users suffering from moderate-to-severe asthma[20] and of younger individuals.

This study also has important strengths. MASK-air® has been developed for patients with rhinitis or asthma and has been assessed in patients with both diseases. VAS asthma — which was the main assessed VAS — has been shown to have high reliability, concurrent validity (with strong correlation with VAS dyspnea,[21] significant correlation with the Asthma Control Test[22] and moderate correlation with CARAT[23]) and moderate responsiveness.[24] We also assessed a sample of participants enrolled by a physician to validate our main results. In addition, this study was conducted in 25 countries (indicating a generalisability of results).

Results were highly consistent when using two clustering methodologies or when assessing different sets of patients. Furthermore, the average number of days reported by patients was longer than in previous MASK-air studies.[20] This longer period of reporting will enable future studies to assess medication adherence.

Interpretation

We classified approximately 70% of the MASK-air® users as having probable asthma or no current asthma (Clusters A, B, C' and D3). In addition, we identified a set of patients who would benefit from further clinical assessment, including users who present high values of VAS Asthma despite not reporting asthma or asthma medications (Clusters C'' and D1). This suggests an under-diagnosis of asthma. Using the Twinning data, most patients of these clusters were classified by their physician as having no asthma. Patients of Cluster A ("uncontrolled treated asthma") may also benefit from clinical assessment for treatment adjustment. It is possible that patients of this cluster may comprise an extreme asthma phenotype, which may be poorly responsive to asthma treatment. Interestingly, this asthma phenotype also tends to display poorer rhinitis control.

Only one-third of the patients with probable asthma reported information that was at least partly compatible with proper treatment/control. This may mirror the clinical challenges related to diagnosing asthma, assessing its severity and tailoring medication. It may also enable patients to understand the importance of self-management.

Some interesting hypothesis-generating results have been observed: (i) There may be an extreme asthma phenotype with a high level of multimorbidity and a relatively poor response to treatment, both for rhinitis and asthma. If this group is confirmed in epidemiologic studies, it may be predictive of the need for biologicals and may allow patient stratification for these treatments. (ii) Better asthma control associated with lower and upper airways as well as eye symptoms. Patients had a similar control for all morbidities, whether or not they received treatment. Ocular symptoms are seldom considered in asthma, although epidemiologic studies have stated their importance.[10] (iii) Among the seven identified clusters, six were associated with asthma and one - rhinitis without current asthma - was strikingly different, suggesting that rhinitis alone and rhinitis and asthma are different diseases.[25]

Taken together, these results suggest that RWD collected under pragmatic circumstances - and particularly when combining information from different variables - can be used to investigate asthma and to identify patients who would benefit from further clinical assessment for diagnostic or therapeutic reasons. This may allow for future studies to be conducted in order to develop CSMSs for the assessment of asthma control based on MASK-air® data.

Conclusion

This study allowed a consistent identification of seven profiles based on the probability of having asthma and on its control. It resulted in a classification supported by physician-diagnosed asthma and in the identification of a substantial percentage of patients potentially benefiting from clinical assessment for diagnosis or treatment adjustment purposes. The use of an mHealth app can help to complement classical epidemiological approaches with RWD. This

can potentially support the identification of patients with asthma and reduce biases of epidemiologic studies solely relying on the retrospective data of self-reported asthma diagnoses.

Data availability

Data are available upon request to Prof. J Bousquet (jean.bousquet@orange.fr).

Funding sources

MASK-air® has been supported by EU grants (POLLAR, EIT Health; Structural and Development Funds, Twinning, EIP on AHA and H2020) and educational grants from Mylan-Viatris, ALK, GSK, Novartis and Uriach. There was no specific funding for this study.

Take-home message

K-means cluster analysis algorithms using real-world data obtained using a mobile app in over 8,000 patients identified patients with probable or possible asthma confirmed by a sub-study in patients with physician-diagnosed asthma.

Conflicts of interest

Dr. Agache has nothing to disclose.

Dr. Amaral has nothing to disclose.

Dr. Antó has nothing to disclose.

Dr. Basagaña has nothing to disclose.

Ms. Bedbrook has nothing to disclose.

Dr. Bergmann has nothing to disclose.

Dr. Bonini has nothing to disclose.

Dr. Bosnic-Anticevich reports grants from TEVA, personal fees from TEVA, personal fees from TEVA, personal fees from AstraZeneca, personal fees from AstraZeneca, personal fees from Boehringer Ingelheim, personal fees from Boehringer Ingelheim, personal fees from GSK, personal fees from Sanofi, personal fees from Mylan, outside the submitted work.

Dr. Boulet reports grants from Amgen, AstraZeneca, GlaxoSmithKline, Merck, Novartis, Sanofi-Regeneron, personal fees from UptoDateÐaylor and FrancisÐptoDateÐaylor and FrancisÐ personal fees from Astra Zeneca, Novartis, GlaxoSmithKline, Merck, Sanofi-Regeneron, personal fees from AstraZeneca, Covis, GlaxoSmithKline, Novartis, Merck, Sanofi, outside the submitted work.

Dr. Bousquet reports personal fees from Chiesi, Cipla, Hikma, Menarini, Mundipharma, Mylan, Novartis, Sanofi-Aventis, Takeda, Teva, Uriach, other from KYomed-Innov, personal fees from Purina, outside the submitted work.

Dr. Brusselle reports personal fees from Astra Zeneca, personal fees from Boehringer-Ingelheim, personal fees from Chiesi, personal fees from GlaxoSmithKline, personal fees from Novartis, personal fees from Sanofi, personal fees from Teva, grants from MerckSharp&Dohme, outside the submitted work.

Dr. Brussino has nothing to disclose.

Dr. Buhl reports grants to Mainz University Hospital from Boehringer Ingelheim, GlaxoSmithKline, Novartis, and Roche, and personal fees from AstraZeneca, Berlin-Chemie, Boehringer Ingelheim, Chiesi, Cipla, GlaxoSmithKline, Novartis, Roche, Sanofi, and Teva, all outside the submitted work.

Dr. Canonica has nothing to disclose.

Dr. Cecchi reports personal fees from Sanofi, personal fees from Astra Zeneca, personal fees from Novartis, personal fees from Thermofisher, personal fees from Menarini, personal fees from Malesci, outside the submitted work.

Dr. Charpin has nothing to disclose.

Dr. Chaves Loureiro has nothing to disclose.

Dr. Cruz reports grants and personal fees from Astrazeneca, personal fees from Boehringer-Ingelheim, personal fees from Chiesi, grants and personal fees from GSK, personal fees from Glenmark, personal fees from Novartis, grants and personal fees from Sanofi, personal fees from Mylan, personal fees from Abdi-Ibrahim, outside the submitted work.

Dr. Czarlewski has nothing to disclose.

Dr. de Blay reports other from Novartis, other from ALK, other from Stallergènes, other from Regeneron, other from DBV, other from Sanofi, other from Boehringer, other from AstraZeneca, outside the submitted work.

Dr. Devillier reports personal fees and non-financial support from Stallergenes Greer, personal fees and non-financial support from ALK-Abello, personal fees and non-financial support from Astra Zeneca, personal fees and non-financial support from CHIESI, personal fees from MENARINI, personal fees and non-financial support from MYLAN / Meda Pharma, personal fees and non-financial support from Novartis, personal fees and non-financial support from GlaxoSmithKline, personal fees and non-financial support from Sanofi, personal fees and non-financial support from IQVIA, outside the submitted work; .

Dr. Fonseca reports participation in SME that has mHealth technologies for patients with asthma.

Dr. Gemicioglu has nothing to disclose.

Dr. Haahtela has nothing to disclose.

Dr. Joos reports grants and personal fees from AstraZeneca, grants and personal fees from Chiesi, personal fees from Bayer, grants and personal fees from GlaxoSmithKline, personal fees from Novartis, personal fees from Lapharcon, personal fees from Eureca vzw, outside the submitted work; all fees were paid to the Department of Respiratory Medicine.

Dr. Jutel reports personal fees from ALK-Abello, personal fees from Allergopharma, personal fees from Stallergenes, personal fees from Anergis, personal fees from Allergy Therapeutics, personal fees from Leti, personal fees from HAL, during the conduct of the study; personal fees from GSK, personal fees from Novartis, personal fees from Teva, personal fees from Takeda, personal fees from Chiesi, outside the submitted work; .

Dr. Klimek has nothing to disclose.

Dr. Kuna reports personal fees from Adamed, personal fees from AstraZeneca, personal fees from Berlin Chemie

Menarini, personal fees from Boehringer Ingelheim, personal fees from Chiesi, personal fees from GSK, personal fees from Novartis, personal fees from Polpharma, outside the submitted work.

Dr. Kupczyk reports personal fees from Astra Zeneca, Novartis, Glaxo, Teva, Adamed, Sanofi, Berlin Chemie, Chiesi, Emma, Lekam, outside the submitted work.

Dr. Kvedariene reports other from Norameda, other from BerlinCHemie Menarini, outside the submitted work.

Dr. Larenas Linnemann reports personal fees from Allakos, Amstrong, Astrazeneca, Chiesi, DBV Technologies, Grunenthal, GSK, Mylan/Viatris, Menarini, MSD, Novartis, Pfizer, Sanofi, Siegfried, UCB, Alakos, Gossamer, Carnot, grants from Sanofi, Astrazeneca, Novartis, Circassia, UCB, GSK, Purina institute, Abvvie, Lilly, Pfizer, outside the submitted work.

Dr. Laune has nothing to disclose.

Dr. Louis reports grants and personal fees from GSK, grants and personal fees from AZ, grants and personal fees from Chiesi, personal fees from Novartis, personal fees from Sanofi, outside the submitted work.

Dr. Pech has nothing to disclose.

Dr. Mäkelä has nothing to disclose.

Dr. Morais-Almeida has nothing to disclose.

Dr. Nadif has nothing to disclose.

Dr. Niedoszytko has nothing to disclose.

Dr. Ohta has nothing to disclose.

Dr. Papadopoulos reports personal fees from Novartis, personal fees from Nutricia, personal fees from HAL, personal fees from MENARINI/FAES FARMA, personal fees from SANOFI, personal fees from MYLAN/MEDA, personal fees from BIOMAY, personal fees from AstraZeneca, personal fees from GSK, personal fees from MSD, personal fees from ASIT BIOTECH, personal fees from Boehringer Ingelheim, grants from Gerolymatos International SA, grants from Capricare, outside the submitted work.

Dr. Papi reports grants from CHIESI, ASTRAZENECA, GSK, BI, PFIZER, TEVA, SANOFI, personal fees from CHIESI, ASTRAZENECA, GSK, NOVARTIS, SANOFI, IQVIA, AVILLION, ELPEN PHARMACEUTICALS, personal fees from CHIESI, ASTRAZENECA, GSK, BI, MENARINI, NOVARTIS, ZAMBON, MUNDIPHARMA, TEVA, SANOFI, EDMOND PHARMA, IQVIA, MSD, AVILLION, ELPEN PHARMACEUTICALS, outside the submitted work.

Dr. Pham-Thi has nothing to disclose.

Dr. Puggioni reports personal fees from SANOFI, personal fees from NOVARTIS, personal fees from ASTRAZENECA, personal fees from GSK, personal fees from VALEAS, personal fees from MENARINI, personal fees from CHIESI, from null, personal fees from STALLERGENS, personal fees from MUNDIPHARMA, outside the submitted work.

Dr. Regateiro has nothing to disclose.

Dr. Rivero Yeverino has nothing to disclose.

Dr. Roche reports grants and personal fees from Boehringer Ingelheim, grants and personal fees from Novartis, grants and personal fees from GSK, personal fees from AstraZeneca, personal fees from Chiesi, grants and personal fees from Pfizer, personal fees from Sanofi, personal fees from Zambon, personal fees from MSD, outside the submitted work.

Dr. Romantowski has nothing to disclose.

Dr. Sá-Sousa has nothing to disclose.

Dr. Samolinski has nothing to disclose.

Dr. Sastre reports grants and personal fees from SANOFI, personal fees from GSK, personal fees from NOVARTIS, personal fees from ASTRA ZENECA, personal fees from MUNDIPHARMA, personal fees from FAES FARMA, outside the submitted work.

Dr. Shamji has nothing to disclose.

Dr. Sheikh reports grants from Asthma UK, outside the submitted work.

Dr. Scichilone has nothing to disclose.

Dr. Sousa-Pinto has nothing to disclose.

Dr. Suppli Ulrik reports grants and personal fees from AZ, persona! fees from GSK, grants and personal fees from BI, personal fees from Chiesi, personal fees from TEVA, personal fees from Orion Pharma, grants and personal fees from Sanofi, grants and personal fees from Novartis, outside the submitted work.

Dr. Taborda-Barata has nothing to disclose.

Dr. Usmani reports grants and personal fees from astra zeneca, grants and personal fees from boehringer ingelheim, grants and personal fees from chiesi, grants and personal fees from glaxosmithkline, personal fees from napp, personal fees from mundipharma, personal fees from sandoz, personal fees from takeda, grants from edmond pharma, personal fees from cipla, personal fees from covis, personal fees from novartis, personal fees from mereo biopharma, personal fees from orion, personal fees from menarini, personal fees from ucb, personal fees from trudell medical, personal fees from deva, personal fees from kamada, personal fees from covis, personal fees from kyorin, outside the submitted work;.

Dr. Valiulis has nothing to disclose.

Dr. Vandenplas has nothing to disclose.

Dr. Ventura has nothing to disclose.

Dr. Yorganc¸oğlu has nothing to disclose.

Dr. Zuberbier: Organizational affiliations: Commitee member: WHO-Initiative "Allergic Rhinitis and Its Impact on Asthma" (ARIA) ; Member of the Board: German Society for Allergy and Clinical Immunology (DGAKI) ; Head: European Centre for Allergy Research Foundation (ECARF) ; President: Global Allergy and Asthma European Network (GA2LEN) ; Member: Committee on Allergy Diagnosis and Molecular Allergology, World Allergy Organization (WAO)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.pulmoe.2022.10.005.

References

1. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211—7.
2. Zhu K, McKnight B, Stergachis A, Daling JR, Levine RS. Comparison of self-report data and medical records data: results from a case-control study on prostate cancer. Int J Epidemiol. 1999;28(3):409—17.

3. Burnett DM, Vardiman JP, Deckert JA, Ward JL, Sharpe MR. Perception of exercise-induced bronchoconstriction in college athletes. Respir Care. 2016;61(7):897—901.
4. Kavanagh J, Jackson DJ, Kent BD. Over- and under-diagnosis in asthma. Breathe. 2019;15(1):e20—7.
5. Aaron SD, Boulet LP, Reddel HK, Gershon AS. Underdiagnosis and overdiagnosis of asthma. Am J Respir Crit Care Med. 2018;198 (8):1012—20.
6. Toren K, Palmqvist M, Lowhagen O, Balder B, Tunsater A. Self-reported asthma was biased in relation to disease severity while reported year of asthma onset was accurate. J Clin Epidemiol. 2006;59(1):90—3.
7. Mirabelli MC, Beavers SF, Flanders WD, Chatterjee AB. Reliability in reporting asthma history and age at asthma onset. J Asthma. 2014;51(9):956—63.
8. Haldar P, Pavord ID, Shaw DE, et al. Cluster analysis and clinical asthma phenotypes. Am J Respir Crit Care Med. 2008;178 (3):218—24.
9. Boudier A, Chanoine S, Accordini S, et al. Data-driven adult asthma phenotypes based on clinical characteristics are associated with asthma outcomes twenty years later. Allergy. 2019;74 (5):953—63.
10. Amaral R, Bousquet J, Pereira AM, et al. Disentangling the heterogeneity of allergic respiratory diseases by latent class analysis reveals novel phenotypes. Allergy. 2019;74(4):698—708.
11. Wu W, Bang S, Bleecker ER, et al. Multiview cluster analysis identifies variable corticosteroid response phenotypes in severe asthma. Am J Respir Crit Care Med. 2019;199(11):1358—67.
12. Tibble H, Chan A, Mitchell EA, et al. A data-driven typology of asthma medication adherence using cluster analysis. Sci Rep. 2020;10(1):14999.
13. Bousquet J, Devillier P, Anto JM, et al. Daily allergic multimorbidity in rhinitis using mobile technology: a novel concept of the MASK study. Allergy. 2018;73(8):1622—31.
14. Menditto E, Costa E, Midao L, et al. Adherence to treatment in allergic rhinitis using mobile technology. The MASK study. Clin Exp Allergy. 2019;49(4):442—60.
15. Bousquet J, Agache I, Aliberti MR, et al. Transfer of innovation on allergic rhinitis and asthma multimorbidity in the elderly (MACVIA-ARIA) - reference site twinning (EIP on AHA). Allergy. 2017;73(1):77—92.
16. Fonseca JA, Nogueira-Silva L, Morais-Almeida M, et al. Control of allergic rhinitis and asthma test (CARAT) can be used to assess individual patients over time. Clin Transl Allergy. 2012;2 (1):16.
17. GINA report 2021. https://ginasthmaorg/wp-content/uploads/2021/05/GINA-Main-Report-2021-V2-WMSpdf. 2021.
18. Laune D, Arnavielhe S, Viart F, et al. [Adaptation of the general data protection regulation (GDPR) to a smartphone app for rhinitis and asthma (MASK-air(R))]. Rev Mal Respir. 2019;36 (9):1019—31.
19. Fonseca JA, Nogueira-Silva L, Morais-Almeida M, et al. Validation of a questionnaire (CARAT10) to assess rhinitis and asthma in patients with asthma. Allergy. 2010;65(8):1042—8.
20. Bedard A, Basagana X, Anto JM, et al. Treatment of allergic rhinitis during and outside the pollen season using mobile technology. A MASK study. Clin Transl Allergy. 2020;10(1): 62.
21. Sousa-Pinto B, Fonseca JA, Gemicioglu B, et al. Patient-reported outcome measures (PROMs) using the MASK-air(R) app in severe asthma. Allergy. 2022;77(5):1600—2.
22. Sastre J, Del Cuvillo A, Colas C, et al. Validation of the MASK-air App for assessment of allergic rhinitis. Allergy. 2020;75 (11):2958—61.
23. Sousa-Pinto B, Sa-Sousa A, Amaral R, et al. Assessment of the control of allergic rhinitis and asthma test (CARAT) using MASK-air. J Allergy Clin Immunol Pract. 2022;10(1):343—5.
24. Sousa-Pinto B, Eklund P, Pfaar O, et al. Validity, reliability, and responsiveness of daily monitoring visual analog scales in MASK-air(R). Clin Transl Allergy. 2021;11(7):e12062.
25. Lemonnier N, Melen E, Jiang Y, et al. A novel whole blood gene expression signature for asthma, dermatitis, and rhinitis multimorbidity in children and adolescents. Allergy. 2020;75:3248—60.