

PAGE Study: Summary of a Study Protocol to Estimate the Prevalence of Severe Asthma in Spain Using Big Data Methods

Almonacid Sánchez C¹, Melero Moreno C², Quirce Gancedo S^{3,8}, Sánchez-Herrero MG⁴, Álvarez Gutiérrez FJ⁵, Bañas Conejero D⁴, Cardona V⁶, Soriano JB^{7,8}

¹Hospital Ramón y Cajal IRYCIS, Madrid, Spain

²Hospital 12 Octubre, Madrid, Spain

³Hospital La Paz, Madrid, Spain

⁴GlaxoSmithKline, Madrid, Spain

⁵Hospital Virgen del Rocío, Sevilla, Spain

⁶Hospital Vall d'Hebron, Barcelona, Spain

⁷Hospital de la Princesa, Madrid, Spain

⁸CIBERES, Spain

J Investig Allergol Clin Immunol 2021; Vol. 31(4): 308-315

doi: 10.18176/jiaci.0483

■ Abstract

Background: The proposal and the initiative for the Prevalence of Severe Asthma in Hospital Units in Spain (PAGE) study came about because of the widespread implementation of electronic medical records and the limited data available on the prevalence of severe asthma in hospitals in our setting.

Objectives: The primary objective was to determine the prevalence of severe asthma in the outpatient clinics of allergy and pulmonology departments in Spain. The secondary objectives were to describe the most prevalent characteristics and phenotypes of severe asthma, to evaluate the selection criteria for receiving approved biological treatments for this disease, and to estimate consumption of resources. Furthermore, digital technology and new data collection sources made it possible to reuse information stored in electronic medical records (Big Data). The study was performed using one such tool, Savana.

Methods: The PAGE study was a multicenter, nonexperimental, observational, cross-sectional study in the first phase and a prospective study in the second phase. The study was controlled and population-based, with 2-stage selection of patients by random sampling. The research was carried out in 40 hospitals selected using convenience sampling in order to ensure geographical representativeness in Spain.

Results: This manuscript describes the study design and protocol.

Conclusions: Our study design was sufficiently robust to avoid bias and to establish the prevalence of patients with severe asthma in Spanish hospitals. It was also the first to incorporate new tools that can help in routine clinical practice and research, such as big data analysis software, and to evaluate the reliability and efficiency of these tools.

Key words: Severe Asthma. Big data. Prevalence. Hospital. Machine learning. Predictions.

■ Resumen

Antecedentes: La propuesta y la iniciativa de realizar el estudio Prevalencia del Asma Grave en las Unidades Hospitalarias de España (PAGE) surge desde la perspectiva de una implantación generalizada de la historia clínica electrónica y de los escasos datos disponibles sobre la prevalencia del asma grave en los hospitales de nuestro entorno.

Objetivos: El objetivo principal es determinar la prevalencia de asma grave en las consultas externas de los servicios de alergia y neumología en España. Como objetivos secundarios, se incluyen describir las características y fenotipos más prevalentes del asma grave, evaluar los criterios de selección para recibir los tratamientos biológicos aprobados para esta enfermedad y estimar el consumo de recursos. Además, aprovechando la incorporación de la tecnología digital y las nuevas fuentes de recogida de datos, que permiten la reutilización de la información almacenada en las historias clínicas electrónicas (Big Data), se ha integrado en el estudio la utilización de una de estas herramientas (Savana).

Métodos: El estudio PAGE está diseñado como un estudio multicéntrico, no experimental, observacional, transversal en una primera fase, y prospectivo en una segunda fase, controlado, basado en la población, con una selección de sujetos en dos etapas por muestreo aleatorio. La investigación se llevará a cabo en 40 hospitales, de acuerdo con un criterio de conveniencia, que asuma la representatividad geográfica de España.

Resultados: El presente manuscrito describe el diseño y protocolo del estudio. Conclusiones: nuestro diseño del estudio es robusto para evitar sesgos y permitir establecer la prevalencia de pacientes con asma grave en los hospitales españoles. Es el primero en incorporar nuevas herramientas que pueden ayudar en la práctica clínica habitual y en la investigación, como un gran software de análisis de datos, y en evaluar su fiabilidad y eficiencia.

Palabras clave: Asma grave. Big data. Prevalencia. Aprendizaje automático. Predicciones.

Introduction

Asthma is one of the most common chronic diseases worldwide. According to the World Health Organization (WHO), more than 358 million people are estimated to have asthma [1,2]. These individuals are heterogeneously distributed throughout the world, irrespective of the degree of development [3].

The prevalence of asthma varies widely even within a single country. In Spain, the mean prevalence of asthma is 5% in the general adult population and 10% in children [4-7]. Existing evidence is more limited regarding the prevalence of severe uncontrolled asthma. The only data available in Spain are from a study conducted in hospital asthma units, where a prevalence of 3.9% was estimated in adults with asthma [8] and 8.8% in children with asthma [7].

According to the recommendations of the Global Initiative for Asthma (GINA) [1], the paradigm for asthma management involves 5 treatment steps, where treatment is adjusted in a continuous cycle based on the degree of control. Current data show that control of asthma is an achievable goal with the available treatment armamentarium [9]. However, analysis of results shows that asthma is not controlled, with different nuances, in a significant proportion of cases [10,11]. The reasons for this include poor adherence to therapy [12], comorbidities, and exposure to aggravating factors [13-15]. This lack of control leads to considerable consumption of resources [16-18], thus making asthma a real public health problem.

Both GINA and the European Respiratory Society (ERS)/American Thoracic Society (ATS) consensus on the definition, evaluation, and treatment of severe asthma define this type of asthma as one that requires treatment with high doses of inhaled corticosteroids (ICS) in addition to maintenance treatment (and/or systemic corticosteroids) (GINA step 4 or 5) to prevent it from being “uncontrolled” or remaining “uncontrolled” despite treatment [16]. This same definition has been adopted by other subsequent guidelines, such as the consensus on severe uncontrolled asthma published by the Spanish Society of Pulmonology and Thoracic Surgery (SEPAR) and adapted to our setting [17]. However, although both documents adequately define the steps for labelling patients as having poorly controlled severe asthma, the reality is that many patients originally classified as having moderate or severe asthma either did not have asthma or had milder manifestations of uncontrolled asthma [17,18].

The Prevalence of Severe Asthma in Hospital Units in Spain (PAGE) study was proposed based on widespread implementation of the electronic medical record and the scant data available on the prevalence of severe asthma in hospitals in our setting. The primary objective was to determine the prevalence of severe asthma in the outpatient clinics of allergy and pulmonology departments in Spain. As secondary objectives, the study sought to describe the characteristics and most prevalent phenotypes of severe asthma, to evaluate the selection criteria for receiving biologic therapies approved for this disease, and to estimate consumption of resources. In addition, taking advantage of the incorporation of digital technology and new data collection sources, which allow

for reuse of information stored in electronic medical records (big data) [19,20], a further objective was to incorporate and use the data management tool, Savana. The term *big data* refers to the processing of large volumes of data using mathematical algorithms in order to establish relationships between data (structured, unstructured, and semistructured) and to determine behavioral patterns that predict trends for improving decision-making. Analysis of big data using traditional methods is excessively time-consuming. In addition, uploading them to a relational database for analysis is expensive. Few studies have analyzed this technology in patients with asthma [19-26]. Savana analyzes and interprets the plain free text contained in electronic medical records, regardless of the electronic system in which these records operate, in order to conduct predictive analyses without the use of classic statistical methods. Savana can also compare

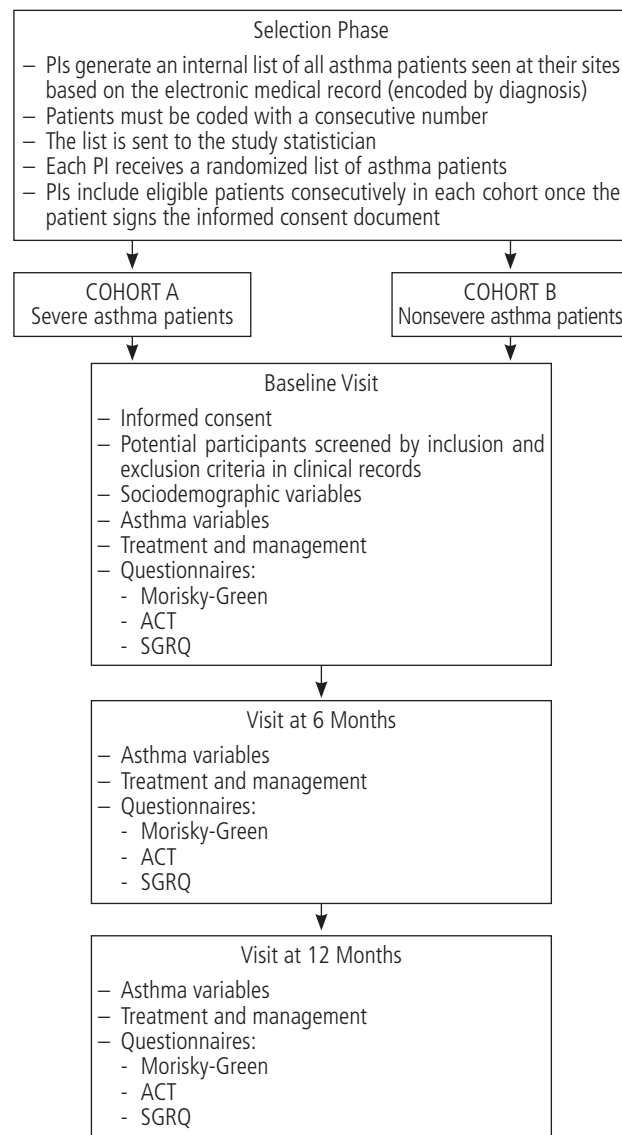


Figure. Study design. PI indicates principal investigator; ACT, Asthma Control Test; SGRQ, Saint George's Respiratory Questionnaire.

records with the descriptive analysis of prevalence and the prospective analysis of the course of patients in the first part of the study, which will be considered the gold standard against which the tool will be compared.

A Scientific Committee consisting of 4 pulmonologists, 2 allergists, and an epidemiologist was formed to prepare the study and provide advice. This study was conducted in accordance with the criteria established in Order SAS/3470/2009 of 16 December, which establishes the guidelines for postauthorization observational studies with medicinal products for human use (Notification/submission to the responsible regulatory authorities, ethics committees, and/or competent authorities) as stipulated by Spanish law (Order SAS/3470/2009 for observational studies). The study was conducted in accordance with the protocol and standard operating procedures to ensure compliance with the International Council on Harmonisation Good Clinical Practice guidelines (1996) [27]. To maintain patient confidentiality, demographic data that could identify the patient (initials, date of birth) were not collected.

Methodology

Design

PAGE was a multicenter, nonexperimental, observational study that was cross-sectional in the first phase and prospective in the second phase. It was a controlled and population-based study in which participants were selected by random sampling in 2 stages. The research was conducted at 40 national hospitals (Figure) selected using convenience sampling in order to ensure geographic representativeness in Spain.

Study Objectives

The study objectives are defined in Table 1.

Selection of Participants

First stage: random population sampling. The principal investigator (PI) at each site created an internal confidential list of all patients diagnosed with asthma at the site through the local diagnostic coding systems. Patients were identified with a consecutive number known only by the investigator, who provided the total number of patients included to the statistician. The statistician then provided a list of 80 random numbers per site. Each investigator, in the order established in the randomized list, consecutively diagnosed the severity of asthma (mild-moderate or severe) in the electronic case report form (eCRF) of 40 patients in order to determine the prevalence of severe asthma.

Second stage: random sampling. Application of randomization to the list of 40 patients in the previous phase led to the selection of 12 patients per site. These were classified into 2 cohorts in a 2 to 1 ratio with regard to severity (severe asthma, nonsevere asthma).

Cohort A. Patients with severe asthma who meet the inclusion criteria and none of the exclusion criteria.

Cohort B. Patients with nonsevere asthma who meet the inclusion criteria and none of the exclusion criteria.

Table 1. PAGE Study Objectives

Primary
To estimate the prevalence of patients with severe asthma seen in the outpatient clinics of allergy and pulmonology departments of Spanish hospitals.
Secondary
<ol style="list-style-type: none"> 1. To describe the sociodemographic and clinical characteristics of patients with severe asthma and to compare them with those of patients with nonsevere asthma (except drug treatment). 2. To estimate the prevalence of the different phenotypes of severe asthma. 3. To describe patients who meet the eligibility criteria for biologic therapy for severe asthma. 4. To assess the course of patients with severe asthma with respect to patients with nonsevere asthma at 6 and 12 months. 5. To assess the impact of severe asthma with respect to that of nonsevere asthma in terms of health-related quality of life. 6. To measure the use of healthcare resources in patients with severe asthma in terms of direct and indirect resources and to compare them with patients with nonsevere asthma. 7. To predict the course of patients with severe and nonsevere asthma at 6 and 12 months, based on the experience and knowledge of the investigator, and to compare the predictions with the results obtained during monitoring (secondary objective 5). 8. To establish the determinants used by clinicians to predict patient outcomes.
Substudy Objectives
The substudy will compare the results obtained using this software installed at 20 sites with the results obtained through the traditional observational study. The tool will be used descriptively (substudy objective 1) and predictively (substudy objectives 2 and 3). The substudy objectives to be measured are as follows: <ol style="list-style-type: none"> 1. To verify whether the prevalence of severe asthma calculated from the aggregate information collected through this specific software is similar to the prevalence of severe asthma obtained in the observational study through monitoring, which will be the gold standard for comparison. 2. To compare the prediction of the course of patients at 6 and 12 months based on collection of aggregate information collected during the 5 years prior to the study with the prediction made by the clinician for the same period. This will only provide descriptive information; no formal comparisons will be made between the prediction made by the clinician and the tool. 3. To compare the prediction of the tool on the course of patients with severe asthma at 6 and 12 months with the results obtained by monitoring in the observational study. The gold standard will be the information collected during monitoring. The tool will be compared to "real data". Further details on secondary objective comparisons can be found in the section <i>Statistical Analysis</i>.

In both cohorts, individuals who refused to participate in the follow-up phase were replaced by the following patient in the randomization list until the required sample size was achieved.

Each patient's data, including demographic data, asthma characteristics, comorbidities, treatment, and laboratory test results, as recorded in the patient's medical record (source document), were recorded in the eCRF. No identifying information was collected; a unique patient number was automatically assigned by the eCRF once the investigator or designated authorized person created the patient file.

Eligibility Criteria

Cohort A: Patients aged ≥ 18 years and diagnosed with severe asthma according to GINA and ERS/ATS criteria who sign the relevant informed consent document.

Cohort B: Patients with nonsevere asthma according to GINA criteria aged ≥ 18 years who sign the relevant informed consent.

Both groups of patients were recruited over a period of 3 months. Subsequent follow-up lasted 12 months according to standard clinical practice (second phase), with information collected at 3 visits (baseline, 6 and 12 months).

Sample Size

In the case of the primary objective, in order to assess the prevalence of patients with severe asthma treated in the allergy and pulmonology units of Spanish hospitals, it was necessary to diagnose 1600 asthmatic patients, considering a maximum inaccuracy estimate of 2% in the prevalence of severe asthma and an asymptotic 2-sided 95%CI and assuming a 20% prevalence of severe asthma and an average of 40 diagnoses per site. A prevalence of severe asthma below the expected value would provide lower values of inaccuracy. For the secondary objectives, a total sample of 320 patients with severe asthma (ie, minimum of 8 patients per site) guaranteed a maximum inaccuracy of 5% in the estimation of the percentages derived from the binary variables, provided that these account for less than 30%. Cohort B includes 160 patients with nonsevere asthma (ie, an average of 4 patients per site).

Data Collection

Primary and secondary data were collected at 3 visits (baseline, 6 and 12 months). The source documents and medical records were used to obtain sociodemographic and clinical data, diagnosis, treatment, and laboratory test results, and data provided directly by the patients during each visit were recorded. Patient-reported outcomes were recorded using paper questionnaires and transcribed by the investigators to an eCRF that was specially designed for the study (<https://ines.emea.research.quintilesims.com/MR/adm/>). The eCRF uses a 128-bit SSL protocol for web communication, which ensures confidentiality between the servers and the investigator's computer and encrypts all the results sent. The eCRF was the source of information from which the study data were analyzed.

The data included in the medical record are protected by the Spanish Personal Data Protection Act [28] and were therefore anonymized at each site by the person responsible for the information before being sent to Savana. In addition, Savana makes it possible to extract unstructured clinical information (natural language or free text) from the electronic clinical records and transform it into reusable and ordered information for research purposes, while maintaining patient anonymity at all times. In addition, the integral clinical contents are detected and scientifically validated using computational linguistic techniques.

Variables and Procedures

Once patients agreed to participate in the study and sign the informed consent document, variables were collected and study procedures performed at 3 follow-up visits (baseline, 6 and 12 months) according to standard clinical practice and with no therapeutic intervention. Data from participating hospitals and study participants were collected.

Sociodemographic and background data were collected at the baseline visit, as follows: age, ethnic group (Caucasian/Latin/African/Asian/other/unknown), educational level, occupational status, housing, smoking, physical examination, body mass index (BMI), diagnosis of respiratory diseases, asthma severity (Table 2), age at onset of asthma, associated

Table 2. GINA Classification of Asthma Severity^a

	Symptoms During the Day	Symptoms at Night	PEF or FEV ₁	Variability in PEF
STEP 1 Intermittent	<1 time a week Asymptomatic and normal PEF between attacks	≤ 2 times a month	$\geq 80\%$	<20%
STEP 2 Mild-persistent	>1 time a week but <1 time a day Attacks may affect activity	>2 times a month	$\geq 80\%$	20-30%
STEP 3 Moderate-persistent	Daily Attacks affect activity	> 1 time a week	60%-80%	>30%
STEP 4 Severe-persistent	Continuous Limited physical activity	Frequent	$\leq 60\%$	>30%

Abbreviations: FEV₁, forced expiratory volume in the first second; PEF, peak expiratory flow.

^aThe presence of one of the features of severity is sufficient to place a patient in that category. Patients at any level of severity—even intermittent asthma—can have severe attacks.

comorbidities, allergies, clinically relevant exacerbations in the previous year and current treatment, score on the asthma control test (ACT) [29], follow-up laboratory tests, spirometry, bronchodilator test, complete blood count, basic biochemistry with IgE levels, and coagulation laboratory parameters. A health-related quality of life questionnaire was administered at each visit, as was the Saint George's Respiratory Questionnaire [30,31], ACT [29], and Morisky-Green treatment adherence questionnaire [32]. Spirometry was performed using a bronchodilator test according to ERS/ATS recommendations [33] based on the reference values for the Mediterranean population [34]. Slow and forced maneuvers were repeated 15-30 minutes after inhalation of 400 µg of albuterol. The result of the bronchodilator test was considered positive when an increase in FVC or FEV₁ greater than 200 mL and greater than 12% from baseline is recorded. Fractional exhaled nitric oxide (FeNO) was measured using the equipment at each site before spirometry and the bronchodilator test following the ERS/ATS guidelines [35].

In addition, data were collected to determine sputum eosinophilia, if performed (otherwise in blood). When available, skin test results and data on use of health care resources were assessed based on the information collected at each study visit over the previous 6 months. The direct costs were as follows: (a) number of scheduled visits (primary care, allergy, pulmonology, other specialist); (b) number of unscheduled visits and reason for the visit; (c) number of emergency department visits related to asthma; (d) number of tests performed owing to severe asthma (pulmonary function, biomarkers, laboratory tests, chest x-ray, computed tomography); (e) hospital admissions to a ward or the intensive care unit for asthma-related reasons in the previous 6 months; (f) referrals to other specialists. Indirect costs comprised the total number of workdays lost due to asthma in the previous 6 months.

Statistical Analysis

Primary objective

Meta-analysis of binary variables (prevalence) according to a fixed-effects model with inverse variance weighting. In the case of heterogeneity, a random effects model was used. In the case of significant heterogeneity and $I^2 > 50\%$, an exploratory post hoc analysis of subgroups or meta regression was performed according to the characteristics of the sites.

Secondary objectives

Univariate and standard bivariate descriptive analysis were performed for categorical, binary, or continuous variables. Relative frequencies and mean 95%CI estimates were obtained. The *t* and Wilcoxon tests were used for comparisons of 2 means, and the *Z* test was performed for comparison of 2 proportions. The χ^2 test was performed to check the independence of qualitative variables, and logistic binary multiple regression models were run to identify differential factors related to the diagnosis of severe asthma. Mixed models were applied for longitudinal data analysis. Descriptive

measures of predictive reliability were also taken (specificity, sensitivity, and positive and negative predictive values), and probability ratios and logistic binary multiple regression models were used to check for relevant factors related to the prediction of clinical events.

The substudy objectives were analyzed by determining whether the point of prevalence estimated by Savana is included in the 95%CI of the prevalence provided by monitoring. Formal comparisons were made between predictions made by the tool and predictions made by the clinician. Descriptive measures of predictive reliability were used (specificity, sensitivity, positive and negative predictive values), as were probability ratios, for the comparison with the clinician's prediction, which is based on the clinical course of patients with severe asthma.

A 5% significance level and 2-sided tests were applied. There were no special considerations for other types of data, and there was no transformation of variables. No new variables were created by transformations, as this was not a confirmatory study, but a descriptive and hypothesis-generating study.

There was no intermediate analysis. Global α adjustments were not expected for multiple comparisons or subgroup analysis.

Implications for PAGE Study Design

The primary objective of the study was to determine the prevalence of severe asthma in allergy and pulmonology units based on population and random sampling. Anonymized data were collected from the 40 participating sites to obtain the prevalence and prediction of the course of patients with severe asthma, which were compared with the results obtained in the prospective study. Savana has developed a free text analysis tool (as natural language) capable of meaningful interpretation of the content included in electronic medical records, regardless of the electronic system in which the records operate, as long as they are available in free text. Given that the program enables the information included in the electronic medical records to be reused, a randomized sample of patients from the site is not required, as all patients with the inclusion criteria will be detected and included in the study. When large numbers of patients are available, more appropriate techniques will be applied to big data processing. All available data were used to generate the model that best represents patient outcome. The population was separated into a training sample and a validation sample to avoid overtraining of the model. The model combined a set of artificial intelligence techniques, which could not be defined in advance but combined approaches ranging from machine learning to deep learning [36-43]. It was possible to use deep and/or recurrent neuronal networks as a guide.

The impact of random errors was minimized, since the system is based on processing large amounts of information (big data). In the clinical area, it is used as a support tool and aids in medical decision-making. The development of this new technology has run in parallel to progressive implementation of the electronic medical record in Spain.

Discussion

Asthma is a common disease affecting all age groups, with a prevalence of approximately 5% to 10% in Western countries [4,7]. However, the prevalence of patients with severe asthma is poorly understood, perhaps because of the limited consensus on a definition of severity in the international community resulting from the paucity of study results. Severity of asthma is known to have a disproportionate effect on both quality of life and treatment costs [44-49].

Epidemiological studies based on postal surveys have shown that 18% to 29% of all patients with asthma have severe disease [50-54]. These estimates are potentially partial, because patients with more severe disease may be more likely to participate in epidemiologic studies involving respiratory disease. Results from previous studies suggest that people with severe asthma represent the largest health economic burden because of the increased risks of hospitalization, visits to the emergency department, loss of lung function, and low quality of life [45,46]. Therefore, it is recommended that this vulnerable group of patients be evaluated and treated by respiratory specialists, although it is unclear whether this recommendation is followed. While asthma is a clinical diagnosis made by physicians based on typical symptoms and clinical examination, many prevalence studies have been conducted with postal or telephone surveys completed by patients or using prescription databases.

In the recent ERS/ATS guidelines [16], the definition of severe asthma was unified in order to adequately classify this type of patients and prevent the heterogeneity of current definitions. In Spain, only 1 study in adults has analyzed the prevalence of severe asthma seen in pulmonology and allergy departments, estimating it to be 3.9% [8]. Available data for the pediatric population show the prevalence of severe asthma to be 8.8% [7]. All these studies have their limitations. They define severity based on clinical criteria and control based on recommendations other than those of GINA, since their design was before 2006, which is when GINA published the criteria for good asthma control. In addition, the definition of severe asthma differed from that proposed by the ERS/ATS consensus, since the study design was prior to it. In our study, patients were selected by searching the computer systems of each hospital based on key words that included the word asthma or the coding of admission diagnoses in the diagnosis section. The search was general, not limited to allergy or pulmonology departments. The selection period covered from the first computer records to the start of the study. Therefore, patient randomization was very robust and avoided selection bias, thus ensuring optimal prevalence values for patients with severe asthma.

The quality of the data obtained in other prevalence studies is also questionable, since most studies involving a very large population sample are based on postal or telephone surveys and therefore do not confirm the diagnosis of asthma using objective measures [10,55]. In the first part of our study, which analyzed the prevalence of patients with severe asthma, we subsequently verified severity in the medical record according to the ERS/ATS consensus criteria [17] in order to define patients with severe asthma. In a second phase, 8 patients with severe asthma and 4 patients with “nonsevere” asthma per site were selected to follow-up and analyze the secondary

objectives of the study. Therefore, our study design can be considered robust because it avoids patient selection bias and the quality of the data provided is high.

Another potential disadvantage of other prevalence studies is that their cross-sectional design means they only provide a snapshot of the different symptoms and signs of severity and their degree of overlap, without determining the dynamics of the disease and the role of aging. However, although patients may enter and exit the “severity phenotypes” defined over time, the overall prevalence of severity in the population is unlikely to change. Therefore, in our study, a sample of the patients initially analyzed was followed up for 1 year.

Study Limitations

Although this is a prospective study with 3 follow-up visits, some of the study variables (eg, use of resources for asthma) were collected by review of the medical record; therefore, the absence of some data may generate information and recall biases that interfere with the results, although most of them can be verified. The limitation of the retrospective part of the study was mitigated in this case, as the sample was randomly selected per site. In addition, there is likely to be some heterogeneity related to how each site determines the list of all its asthmatic patients and the source (database) from which this list comes. Furthermore, patients may have forgotten to complete the health-related quality of life questionnaires or the ACT, thus potentially interfering with the results related to these variables. However, our study is the first in Spain to examine the contributions of large-scale asthma data analysis (big data). This section is part of a substudy to analyze the reliability and feasibility of big data using as a comparator (gold standard), namely, a conventional analysis of the patient sample followed for 1 year. The substudy will be performed using a computer program specifically designed for medical data analysis called Savana. Twenty of the participating sites will install this software, which performs a comprehensive analysis of all medical records entered into the electronic medical record system. Based on search criteria, Savana is intended to be able to provide the prevalence of patients with asthma, as well as the prevalence by severity, and its results are correlated with the secondary objectives of the study.

In conclusion, our study design proved to be sufficiently robust in order to avoid bias and establish the prevalence of patients with severe asthma in Spanish hospitals. It is the first to incorporate new tools, such as big data analysis software, that may help in routine clinical practice and research and to evaluate the reliability and efficiency of these tools.

Funding

This study was sponsored by GSK (205807).

Conflicts of Interest

MGSB and DBC are employees of GSK.

CAS participated in speaking activities and advisory boards and provided consultancy services during the period 2015-2019 sponsored by AstraZeneca, Boehringer-Ingelheim, Chiesi, GSK, ALK Mundipharma, Novartis, Pfizer, SEPAR,

and NEUMOMADRID. CAS declares not having received, directly or indirectly, funding from the tobacco industry or its affiliates.

JBS participated in speaking activities and advisory boards and provided consultancy services during the period 2015-2019 for Almirall, AstraZeneca, Boehringer-Ingelheim, CHEST, Chiesi, ERS, GEBRO, Grifols, GSK, Linde, Lipopharma, Mundipharma, Novartis, Pfizer, RiRL, Rovi, Sandoz, SEPAR, and Takeda. JBS declares not having received, directly or indirectly, funding from the tobacco industry or its affiliates.

The remaining authors declare that they have no conflicts of interest.

Previous Presentation

XXX Congreso de la Sociedad Española de Alergología e Inmunología Clínica, 24-27 October 2018, Valencia, Spain.

References

- GINA Report 2017: Global Strategy for Asthma Management and Prevention. Global Initiative for Asthma - GINA, <http://ginasthma.org/2017-gina-report-global-strategy-for-asthma-management-and-prevention/> (Accessed 2 September 2017).
- GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med*. 2017;5:691-706.
- Enfermedades respiratorias crónicas. Asma. Organización Mundial de la Salud [Accessed 02 February 2017]. Available at: <http://www.who.int/respiratory/asthma/es/>.
- Guía Española para el Manejo del Asma. GEMA4.2 [Internet]. [Accessed 6 August 2017]. Available at: <http://www.gemasma.com/>.
- European study on asthma. Prevalence of bronchial hyperreactivity and asthma in young adults from 5 Spanish areas. Spanish Group of the European Study on Asthma. *Med Clin*. 1996;106:761-7.
- Arnedo A, Bellido JB, Pac MR, Artero A, Campos JB, Museros L, et al. Incidence of asthma and risk factors in a cohort of schoolchildren aged from 6-7 years old to 14-15 years old in Castellón (Spain) following the International Study of Asthma and Allergies in Childhood (ISAAC). *Med Clin (Barc)*. 2007;129:165-70.
- Plaza-Martín AM, Vennera MC, Galera J, Herráez L, PREX Study Group. Prevalence and clinical profile of difficult-to-control severe asthma in children: results from pneumology and allergy hospital units in Spain. *Allergol Immunopathol*. 2014;42:510-7.
- Quirce S, Plaza V, Picado C, Vennera M, Casafont J. Prevalence of uncontrolled severe persistent asthma in pneumology and allergy hospital units in Spain. *J Investig Allergol Clin Immunol*. 2011;21:466-71.
- Bateman ED, Boushey HA, Bousquet J, Busse WW, Clark TJ, Pauwels RA, et al. Can guideline-defined asthma control be achieved? The Gaining Optimal Asthma Control study. *Am J Respir Crit Care Med*. 2004;170:836-44.
- Demoly P, Annunziata K, Gubba E, Adamek L. Repeated cross-sectional survey of patient-reported asthma control in Europe in the past 5 years. *Eur Respir Rev*. 2012;21:66-74.
- Braido F, Brusselle G, Guastalla D, Ingrassia E, Nicolini G, Price D, et al. Determinants and impact of suboptimal asthma control in Europe: The INTERNATIONAL CROSS-SECTIONAL AND LONGITUDINAL ASSESSMENT ON ASTHMA CONTROL (LIAISON) study. *Respir Res*. 2016;17:51.
- Boulet LP, Vervloet D, Magar Y, Foster JM. Adherence: the goal to control asthma. *Clin Chest Med*. 2012;33:405-17.
- Ledford DK, Lockey RF. Asthma and comorbidities. *Curr Opin Allergy Clin Immunol*. 2013;13:78-86.
- Su X, Ren Y, Li M, Zhao X, Kong L, Kang J. Prevalence of Comorbidities in Asthma and Nonasthma Patients: A Meta-analysis. *Medicine*. 2016;95:e3459.
- Cazzola M, Segreti A, Calzetta L, Rogliani P. Comorbidities of asthma: current knowledge and future research needs. *Curr Opin Pulm Med*. 2013;19:36-41.
- Chung KF, Wenzel SE, Brozek JL, Bush A, Castro M, Sterk PJ, et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur Respir J*. 2014;43:343-73.
- Cisneros Serrano C, Melero Moreno C, Almonacid Sánchez C, Perpiñá Tordera M, Picado Valles C, Martínez Moragón E, et al. Guidelines for severe uncontrolled asthma. *Arch Bronconeumol*. 2015;51:235-46.
- Chen Y, Elenee Argentinis JD, Weber G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clin Ther*. 2016;38:688-701.
- Popovic JR. Distributed data networks: a blueprint for Big Data sharing and healthcare analytics. *Ann N Y Acad Sci*. 2017;1387:105-11.
- Dimitrov DV. Medical Internet of Things and Big Data in Healthcare. *Healthc Inform Res*. 2016;22:156-63.
- Belgrave D, Henderson J, Simpson A, Buchan I, Bishop C, Custovic A. Disaggregating asthma: Big investigation versus big data. *J Allergy Clin Immunol*. 2017;139:400-7.
- Kuo CY, Pan RH, Chan CK, Wu CY, Phan DV, Chan CL. Application of a Time-Stratified Case-Crossover Design to Explore the Effects of Air Pollution and Season on Childhood Asthma Hospitalization in Cities of Differing Urban Patterns: Big Data Analytics of Government Open Data. *Int J Environ Res Public Health*. 2018;15:pii E647.
- Souliotis K, Kani C, Papageorgiou M, Lionis D, Gourgoulanis K. Using Big Data to Assess Prescribing Patterns in Greece: The Case of Chronic Obstructive Pulmonary Disease. *PloS One*. 2016;11:e0154960.
- Stelmach R, Cruz AA. The paradox of asthma: neglect, burden, and big data. *J Bras Pneumol*. 2017;43:159-60.
- Diver S, Brightling CE. Big asthma data: getting bigger and more beautiful? *Thorax*. 2018;73:311-2.
- Ram S, Zhang W, Williams M, Pengetnze Y. Predicting asthma-related emergency department visits using big data. *IEEE J Biomed Health Inform*. 2015;19:1216-23.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Guideline for good clinical practice. Geneva: ICH Harmonised Tripartite Guideline; 1996.
- Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. Available at: <https://www.boe.es/buscar/act.php?id=BOE-A-1999-23750>.

29. Vega JM, Badia X, Badiola C, López-Viña A, Olaguibel JM, Picado C, et al. Validation of the Spanish version of the Asthma Control Test (ACT). *J Asthma*. 2007;44:867-72.
30. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis*. 1992;145:1321-7.
31. Sanjuás C, Alonso J, Prieto L, Ferrer M, Broquetas JM, Antó JM. Health-related quality of life in asthma: a comparison between the St George's Respiratory Questionnaire and the Asthma Quality of Life Questionnaire. *Qual Life Res*. 2002;11:729-38.
32. Morisky DE, DiMatteo MR. Improving the measurement of self-reported medication nonadherence: Final response. *J Clin Epidemiol*. 2011;64:258-63.
33. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26:319-38.
34. García-Río F, Calle M, Burgos F, Casan P, Del Campo F, Galdiz JB, et al. Spirometry. Spanish Society of Pulmonology and Thoracic Surgery (SEPAR). *Arch Bronconeumol*. 2013;49:388-401.
35. Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. *Am J Respir Crit Care Med*. 2011;184:602-15.
36. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132:1920-30.
37. Giger ML. Machine Learning in Medical Imaging. *J Am Coll Radiol*. 2018;15:512-20.
38. Ho LV, Ledbetter D, Aczon M, Wetzel R. The Dependence of Machine Learning on Electronic Medical Record Quality. *AMIA Annu Symp Proc*. 2018;2017:883-91.
39. Shenoy VN, Aalami OO. Utilizing Smartphone-Based Machine Learning in Medical Monitor Data Collection: Seven Segment Digit Recognition. *AMIA Annu Symp Proc*. 2018;2017:1564-70.
40. Choi H, Na KJ. A Risk Stratification Model for Lung Cancer Based on Gene Coexpression Network and Deep Learning. *Biomed Res Int*. 2018;2018:2914280.
41. Hua KL, Hsu CH, Hidayati SC, Cheng WH, Chen YJ. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco Targets Ther*. 2015;8:2015-22.
42. Klang E. Deep learning and medical imaging. *J Thorac Dis*. 2018;10:1325-8.
43. Song Q, Zhao L, Luo X, Dou X. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. *J Healthc Eng*. 2017;2017:8314740.
44. Busse WW, Banks-Schlegel S, Wenzel SE. Pathophysiology of severe asthma. *J Allergy Clin Immunol*. 2000;106:1033-42.
45. Barnes PJ, Woolcock AJ. Difficult asthma. *Eur Respir J*. 1998;12:1209-18.
46. The ENFUMOSA cross-sectional European multicentre study of the clinical phenotype of chronic severe asthma. European Respiratory Society, <http://erj.ersjournals.com/content/22/3/470> (Accessed 29 December 2017).
47. Nolte H, Nepper-Christensen S, Backer V. Unawareness and undertreatment of asthma and allergic rhinitis in a general population. *Respir Med*. 2006;100:354-62.
48. Zuberbier T, Lötvalld J, Simoons S, Subramanian SV, Church MK. Economic burden of inadequate management of allergic diseases in the European Union: a GA2LEN review. *Allergy*. 2014;69:1275-9.
49. Martínez-Moragón E, Serra-Batlles J, De Diego A, Palop M, Casan P, Rubio-Terrés C, et al. Economic cost of treating the patient with asthma in Spain: the AsmaCost study. *Arch Bronconeumol*. 2009;45:481-6.
50. Ekerljung L, Bossios A, Lötvalld J, Olin AC, Rönmark E, Wennergren G, et al. Multi-symptom asthma as an indication of disease severity in epidemiology. *Eur Respir J*. 2011;38:825-32.
51. Cazzoletti L, Marcon A, Corsico A, Janson C, Jarvis D, Pin I, et al. Asthma severity according to Global Initiative for Asthma and its determinants: an international study. *Int Arch Allergy Immunol*. 2010;151:70-9.
52. de Marco R, Marcon A, Jarvis D, Accordini S, Almar E, Bugiani M, et al. Prognostic factors of asthma severity: A 9-year international prospective cohort study. *J Allergy Clin Immunol*. 2006;117:1249-56.
53. von Bülow A, Kriegbaum M, Backer V, Porsbjerg C. The prevalence of severe asthma and low asthma control among Danish adults. *J Allergy Clin Immunol Pract*. 2014;2:759-67.
54. Rönmark E, Lindberg A, Watson L, Lundbäck B. Outcome and severity of adult onset asthma—Report from the obstructive lung disease in northern Sweden studies (OLIN). *Respir Med*. 2007;101:2370-7.
55. Rabe KF, Adachi M, Lai CK, Soriano JB, Vermeire PA, Weiss KB, et al. Worldwide severity and control of asthma in children and adults: the global asthma insights and reality surveys. *J Allergy Clin Immunol*. 2004;114:40-7.

■ *Manuscript received October 21, 2019; accepted for publication January 8, 2020.*

■ **Carlos Almonacid Sánchez**

Hospital Universitario Ramón y Cajal
Ctra. de Colmenar Viejo km. 9,100
28034 Madrid
E-mail: carlos.almonacid@separ.es