

VNIVERSITAT E VALÈNCIA

FACULTAD DE MEDICINA Y ODONTOLOGÍA

DEPARTAMENTO DE MEDICINA PREVENTIVA Y SALUD PÚBLICA, CIENCIAS DE LA
ALIMENTACIÓN, TOXICOLOGÍA Y MEDICINA LEGAL.

Programa de Doctorado 3139, Medicina

Línea de investigación: Salud Pública, Higiene y Sanidad Ambiental.



Epidemiología de la alergia a nivel nacional y por comunidades autónomas. Estudio de prevalencia y factores asociados en un periodo de 25 años

TESIS DOCTORAL

Presentada por: **Marta Burches Feliciano**

Graduada en Medicina

Dirigida por: **Dra. Carmen Saiz Sánchez**

Valencia 2023



VNIVERSITAT
DE VALÈNCIA

Dña Carmen Saiz Sánchez, Doctora en Medicina y Profesora titular de la Universidad de Valencia del Departamento de Medicina Preventiva y Salud Pública, Ciencias de la Alimentación, Toxicología y Medicina Legal y directora de esta tesis.

CERTIFICA:

Que Marta Burches Feliciano, Graduada en Medicina por la Universidad Cardenal Herrera-CEU de Castellón, ha realizado bajo mi dirección el presente trabajo de investigación, titulado "Epidemiología de la alergia a nivel nacional y por comunidades autónomas. Estudio de prevalencia y factores asociados en un periodo de 25 años", el cual posee la suficiente calidad científica para ser presentado para la obtención del Grado de Doctor por la Universidad de Valencia.

Y para que así conste, y a los efectos oportunos, se expide el presente certificado en Valencia a 11 de mayo de 2023.

Fdo. Dra. Carmen Saiz Sánchez

Agradecimientos

En primer lugar me gustaría expresar mi agradecimiento a mi directora de tesis, Dra Carmen Saiz, una persona de excepcional generosidad que me ha guiado durante este proceso con su experiencia investigadora y aportándome confianza, cariño, dedicación, disponibilidad y ayuda en todo momento. Gracias a sus consejos que han sido fundamentales para llevar a cabo y finalizar este trabajo.

Agradezco de manera especial a mi padre, persona que admiro personal y profesionalmente, por su confianza, amor y apoyo incondicional, por su paciencia y motivación, por estar siempre cerca y ser un ejemplo de entrega y bondad. Así como, por haberme iniciado en la especialidad que compartimos y estimular mi interés al conocimiento científico.

Dedicatoria

Con todo mi corazón:

A mi madre y a mi padre. A mi familia.

*A mi marido Quique y a nuestra hija que estamos esperando con gran ilusión,
mis dos soles.*

RESUMEN

Actualmente, se percibe que la alergia se comporta como una patología que se extiende cada vez más, afectando a una proporción importante de la población, y que aparece fundamentalmente en países industrializados con modos de vida occidentales, siendo factores impulsores la contaminación y los efectos del cambio climático.

Sin embargo, cuando acudimos a la búsqueda de confirmación de estos hechos, nos encontramos que los datos epidemiológicos de alergia y asma en nuestro país son muy escasos. Existen algunos estudios de prevalencia de alergia, pero realizados únicamente en población infantil. También algunos muestreos sobre pacientes que ya acudían al alergólogo con la sospecha de una enfermedad de origen alérgico.

Por lo que conocemos, este es el primer estudio que utiliza datos agregados acerca del diagnóstico de alergia y asma, recogidos por instituciones nacionales como son el Ministerio de Sanidad y el Instituto Nacional de Estadística. Se ha realizado un estudio observacional en una ventana o serie temporal de 25 años, recogiendo el diagnóstico de alergia y asma junto con factores sociodemográficos como edad, sexo, tipo de población, nivel de estudios, nivel económico, ocupación, etc. Las herramientas empleadas han sido de tipo estadístico descriptivo y análisis multivariante, éste último tampoco aplicado nunca al tema del estudio.

Los resultados a destacar son, en primer lugar, la confirmación de que el aumento de prevalencia de la alergia en 25 años en España es muy importante (59,09%). En todos los grupos de edad y sexo el aumento ha sido constante en el tiempo, pero sobre todo en los más jóvenes y en las mujeres. Las cifras más altas coinciden con el grupo de personas con estudios superiores y con los niveles de renta más altos. Por el contrario, el asma no ha sufrido la misma evolución y sus cifras se han estabilizado con los máximos en los grupos de más edad. En los grandes núcleos de población aumenta la prevalencia de alergia (50,18%) pero no de asma. Las cifras de alergia y asma en la Comunidad Valenciana siguen el patrón común pero con menor proporción que otras comunidades. El análisis multivariante en su versión estadística y su aplicación gráfica sugiere asociaciones con otros factores de salud como el sedentarismo.

Palabras clave: epidemiología, prevalencia, alergia, asma, análisis multivariante.

ABSTRACT

Currently, it is perceived that allergy behaves like a pathology that is spreading more and more, affecting a significant proportion of the population, and that appears mainly in industrialized countries with western lifestyles, pollution and the effects of climate change.

However, when we look for confirmation of these facts, we find that the epidemiological data of allergy and asthma in our country are very scarce. There are some studies of allergy prevalence, but only performed in children. Also some samples about patients who already went to the allergist with the suspicion of a disease of allergic origin.

As far as we know, this is the first study using aggregated data on allergy and asthma diagnosis, collected by national institutions such as the Ministry of Health and the National Institute of Statistics. An observational study has been conducted in a window or time series of 25 years, collecting the diagnosis of allergy and asthma along with sociodemographic factors such as age, sex, population type, level of education, economic level, occupation, etc. The tools used have been descriptive statistics and multivariate analysis, the latter never applied to the subject of the study.

The results to be highlighted are, first, the confirmation that the increase in allergy prevalence in 25 years in Spain is very important (59,09%). In all age groups and sex the increase has been constant over time, but especially in the younger ones and women. The highest figures agree with the group of people with higher education and with the highest income levels. In contrast, asthma has not undergone the same evolution and its figures have stabilized with the maxima in the older groups. In large population centers, the prevalence of allergy increases (50,18%) but not of asthma. Allergy and asthma figures in the Valencia region follow the common pattern but with a lower proportion than other communities. The multivariate analysis in its statistical version and its graphical application suggest associations with other health factors such as sedentarism.

Key words: epidemiology, prevalence, allergy, asthma, multivariate analysis.

ÍNDICE

I. INTRODUCCION.....	31
1.1. Resumen histórico de la alergia.....	31
1.2. Epidemiología.....	46
1.2.1. Clasificación y tipos de estudios epidemiológicos.....	48
1.2.1.1. Estudios descriptivos.....	48
1.2.1.2. Estudios analíticos.....	51
1.2.1.3. Estudios experimentales.....	51
1.3. Estudios ecológicos.....	52
1.3.1. Clasificación de los estudios ecológicos.....	56
1.4. Epidemiología de la alergia.....	57
1.5. Datos no estructurados: un mundo por explorar, Big Data.....	73
1.5.1. Importancia de la cantidad de datos producidos a nivel mundial.....	73
1.5.2. Tipos de datos.....	79
1.5.3. Forma de analizar los datos.....	82
1.5.3.1. Clasificación de las técnicas estadísticas multivariadas.....	84
1.5.3.1.1. Métodos explicativos o de dependencia.....	85
1.5.3.1.2. Métodos descriptivos o de interdependencia.....	87
1.6. Encuesta Nacional de Salud.....	92
1.6.1. Generalidades.....	92
1.6.2. Objetivos de las Encuestas Nacionales de Salud.....	95
1.6.3. Contenido de las Encuestas Nacionales de Salud.....	96
1.6.4. Características fundamentales objeto de estudio	98
1.6.4.1. Morbilidad percibida por la población	98
1.6.4.2. Características y condiciones de uso de los servicios de salud.....	98
1.6.4.3. Hábitos de vida.....	98
1.6.4.4. Características específicas de la población infantil.....	99
1.6.5. Fases de la encuesta.....	99
1.6.6. Ámbito de la encuesta.....	100
1.6.5.1. Ámbito poblacional.....	100
1.6.5.2. Ámbito geográfico.....	100

1.6.5.3. Ámbito temporal.....	101
1.6.7. Diseño de la muestra.....	101
1.6.7.1. Tipo de muestreo	101
1.6.7.2. Tamaño de la muestra.....	102
1.6.7.3. Metodología estadística.....	102
1.7. Justificación.....	103
II. HIPÓTESIS Y OBJETIVOS.....	105
2.1 Hipótesis.....	105
2.2.1. Objetivo general.....	105
2.2.2. Objetivos específicos.....	105
III. MATERIAL Y MÉTODO.....	107
3.1. Tipo de estudio y descripción de la estadística empleada.....	107
3.2. Tipos de variables.....	108
3.3. Tipos de parámetros estadísticos.....	109
3.3.1. Medidas de centralización.....	109
3.3.2. Medidas de posición.....	109
3.3.3. Medidas de distribución	110
3.3.4. Medidas de dispersion.....	110
3.3.5. Medidas de asociación.....	111
3.3.6. Análisis multivariante.....	112
3.4. Fuentes de datos: La Encuesta Nacional de Salud.....	115
3.5. Grupos de datos incluidos en la ENS empleados en el presente estudio.....	119
3.6. Estadística descriptiva realizada en este estudio.....	120
3.7. Gráficos y su descripción.....	123
3.8. Estadística para verificar la normalidad y detectar la significatividad..	127
3.9. Análisis de conglomerados y de componentes principales.....	128
3.9.1. Análisis de conglomerados.....	129
3.9.2. Análisis de componentes principales.....	131
3.10. Análisis de actividad física y relación con alergia y asma.....	134
3.11. Análisis de factores de salud y relación con alergia y asma.....	135
3.12. Comparativa entre comunidades autónomas de la prevalencia de alergia y asma en 2011 y 2017.....	136

IV. RESULTADOS.....	139
4.1. Estadística descriptiva.....	139
4.1.1. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años.....	139
4.1.2. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años con respecto al tipo de población.....	150
4.1.3. Estadística descriptiva de la prevalencia de alergia según nivel de estudios.....	160
4.1.4. Estadística descriptiva de la prevalencia de alergia según nivel económico.....	165
4.1.5. Estadística descriptiva de la prevalencia de alergia según tipo de ocupación.....	174
4.1.6. Estadística descriptiva de la prevalencia de alergia según tipo de jornada laboral.....	179
4.1.7. Estadística descriptiva de la prevalencia de asma.....	183
4.1.8. Estadística descriptiva de la prevalencia de asma según tipo de población.....	191
4.1.9. Estadística descriptiva de la prevalencia de asma según nivel de estudios.....	199
4.1.10. Estadística descriptiva de la prevalencia de asma según nivel económico.....	204
4.1.11. Estadística descriptiva de la prevalencia de asma según tipo de ocupación.....	210
4.1.12. Estadística descriptiva de la prevalencia de asma según tipo de jornada laboral.....	214
4.2. Conglomerados o <i>clusters</i>	218
4.2.1. Análisis de conglomerados de la evolución de la prevalencia de alergia en 25 años.....	218
4.2.2. Análisis de conglomerados de la evolución de la prevalencia de alergia según tipo de población.....	220
4.2.3. Análisis de conglomerados de la evolución de la prevalencia de alergia según nivel de estudios.....	223

4.2.4. Análisis de conglomerados de la evolución de la prevalencia de asma.....	226
4.2.5. Análisis de conglomerados de la evolución de la prevalencia de asma según tipo de población.....	227
4.2.6. Análisis de conglomerados de la evolución de la prevalencia de asma según nivel de estudios.....	230
4.3. Análisis de componentes principales.....	233
4.3.1. Análisis de componentes principales de la evolución de la prevalencia de alergia en 25 años	234
4.3.2. Análisis de componentes principales de la evolución de la prevalencia de alergia según tipo de población.....	236
4.3.3. Análisis de componentes principales. de la evolución de la prevalencia de alergia según nivel de estudios.....	238
4.3.4. Análisis de componentes principales de la evolución de la prevalencia de asma en 25 años	241
4.3.5. Análisis de componentes principales de la evolución de la prevalencia de asma según el tipo de población.....	243
4.3.6. Análisis de componentes principales de la prevalencia de asma según nivel de estudios.....	245
4.4. Actividad física en relación con alergia y asma	248
4.4.1. Actividad física en relación con alergia y asma de 2011 en comunidades autónomas.....	248
4.4.1.1. Estadística descriptiva de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas.....	250
4.4.1.2. Matriz de correlaciones de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas.....	251
4.4.1.3. Análisis de componentes principales de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas....	252
4.4.1.4. Análisis de conglomerados de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas.....	255
4.4.2. Actividad física en relación con alergia y asma de 2017 en comunidades autónomas.....	260

4.4.2.1. Estadística descriptiva de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas.....	260
4.4.2.2. Matriz de correlaciones de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas.....	262
4.4.2.3. Análisis de componentes principales de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas....	264
4.4.2.4. Análisis de conglomerados de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas.....	267
4.5. Factores de salud y alergia-asma en comunidades autónomas.....	270
4.5.1. Estadística descriptiva de factores de salud en relación con alergia y asma en comunidades autónomas.....	274
4.5.2. Matriz de correlaciones de factores de salud en relación con alergia y asma en comunidades autónomas.....	272
4.5.3. Análisis de componentes principales de factores de salud en relación con alergia y asma en comunidades autónomas.....	274
4.5.4. Análisis de conglomerados de factores de salud en relación con alergia y asma en comunidades autónomas.....	278
4.6 Comparativa de la prevalencia de alergia y asma por comunidades autónomas.....	286
V. DISCUSIÓN.....	299
VI. CONCLUSIONES.....	323
VII.BIBLIOGRAFIA.....	327
VIII. ANEXOS.....	351
8.1. Certificado del comité de ética.....	351

LISTA DE ABREVIATURAS

Las abreviaturas utilizadas en esta tesis son las siguientes:

AFC	Análisis factorial común
ANCOVA	Análisis de covarianza
ANOVA	Análisis de varianza
BCG	Bacilo de Calmette-Guérin
CAPI	Entrevista personal asistida por ordenador
CATI	<i>Computer-assisted Telephone Interview technique</i>
CIS	Centro de Investigaciones Sociológicas
CNED	Clasificación Nacional de Educación
CO2	Dióxido de carbono
CRAN	<i>Comprehensive R Archive Network</i>
DTP	Difteria, tétanos, tos ferina
EAACI	Academia Europa de Alergología e Inmunología Clínica
EESE	Encuesta Europea de Salud en España
EHIS	<i>European Health Interview Survey</i>
ENS	Encuesta Nacional de Salud
GAN	<i>Global Asthma Network</i>
GNU	<i>General Public Licence</i>
H ₀	Hipótesis nula
H ₁	Hipótesis alternativa
IC	Intervalos de confianza
IgE	Inmunoglobulina E específica
IgG3, IgG4	Subclase 3 y 4 de la inmunoglobulina G
IL	Interleucina
IMC	Índice de Masa Corporal
INE	Instituto Nacional de Estadística
ISAAC	<i>International Study of Asthma and Allergies in Childhood</i>
MANCOVA	Análisis multivariado de covarianza
MANOVA	Análisis multivariado de varianza
MCH	Complejo principal de histocompatibilidad

NO2	Óxidos de nitrógeno
OMS/WHO	Organización Mundial de la Salud/ <i>World Health Organization</i>
PIB	Producto Interior Bruto
PM10	Material particulado. El número indica el tamaño
PNB	Producto Nacional Bruto
RDA	República Democrática Alemana
SEE	Sociedad Española de Epidemiología
SO2	Dióxido de azufre

LISTA DE FIGURAS

FIGURA 1. Diagrama de barras con el porcentaje de prevalencia en varones en ordenadas y en abcisas los diferentes grupos según el año de la encuesta con subdivisiones de colores distintos (relación de grupo de edad y sexo) con su significado en la leyenda del pie de la figura.....	141
FIGURA 2. Diagrama de barras con el porcentaje de prevalencia en mujeres en ordenadas y en abcisas los diferentes grupos según el año de la encuesta con subdivisiones de colores distintos (relación de grupo de edad y sexo) con su significado en la leyenda del pie de la figura.....	142
FIGURA 3. Representación gráfica en forma de picos montañosos del porcentaje de prevalencia de alergia diferenciándose por edad y sexo, y distinguiendo los años de medición. En la base de la figura la leyenda explica el porcentaje en forma de diferentes colores.....	143
FIGURA 4. Diagrama de barras acumulado representando en ordenadas el porcentaje de prevalencia de alergia y en abcisas, diferenciando por sexos, la distribución por grupos de edad. Los años de medición representados por diferentes colores en la leyenda.....	144
FIGURA 5. Serie temporal de prevalencia de alergia desde 1993 a 2017. En ordenadas se expresa el porcentaje de prevalencia y en abcisas cada uno de los años en que se efectuaron las encuestas. Numéricamente también se expresa la prevalencia sobre la gráfica y al pie como leyenda.....	145
FIGURA 6. Biplot representando una serie de ejes que son los años de medición en que se efectúan las observaciones, incorporando en ellos las cifras de prevalencia. Proyectándose sobre los ejes cada una de las observaciones numeradas que son los grupos de edad y sexo.....	149
FIGURA 7. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de población según número de habitantes con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	151
FIGURA 8. Serie temporal de la evolución de prevalencia de alergia en 25 años (en la leyenda a pie de tabla se explica el color de cada año). En abcisas el tipo de población y con 3 divisiones: as –cifra total ambos sexos-, v-varones-, m-mujeres-. En ordenadas el porcentaje de prevalencia.....	154
FIGURA 9. Serie temporal de 25 años representando en abcisas las series de años y en ordenadas el porcentaje de prevalencia de alergia en cifras totales, caracterizando por colores el tipo de población.....	155
FIGURA 10. Serie temporal de la evolución de prevalencia de alergia en sucesivos años (en la leyenda a pie de tabla se explica el color que corresponde a cada año) teniendo en abcisas el tipo de población y en ordenadas el porcentaje de prevalencia.....	156
FIGURA 11. Biplot de prevalencia de alergia utilizando como ejes a los años de medición (numerados por orden, desde el 1 al 9, desde 1993 a 2017) situándose sobre los ejes los porcentajes de prevalencia y proyectándose en los ejes cada uno de los nichos de población y sexo (cuadrados rojos que representan cada una de las mediciones desde 1993 a 2017).....	157
FIGURA 12. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de niveles de estudios con 3 divisiones: cifras totales en ambos sexos, varones y	

mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	161
FIGURA 13. Promedio de la prevalencia de alergia en 12 años, expresando en ordenadas el porcentaje de prevalencia y en abcisas los distintos grupos en función del nivel de estudios y diferenciando en cifras totales y sexo.....	162
FIGURA 14. Biplot de prevalencia de alergia según estudios representando en los ejes los niveles de estudios (sin estudios, 1º, 2º, 2º superior, 3º) y en eje diagonal los años de medición. La variables representan los grupos por sexo y nivel de estudios.....	164
FIGURA 15. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes niveles económicos con 3 divisiones: cifras totales en ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	167
FIGURA 16. Promedio de evolución de prevalencia de alergia en 17 años, representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos niveles económicos diferenciándose por cifras totales y sexo.....	169
FIGURA 17. Serie temporal en 15 años de prevalencia de alergia representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos niveles económicos diferenciándose por cifras totales de ambos sexos y desglosados en varones y mujeres. Con desglose según el año de medición que se expresa en diferentes colores según leyenda de pie de figura.....	170
FIGURA 18. Serie temporal en 15 años de prevalencia de alergia en varones representando el porcentaje de prevalencia en ordenadas y en abcisas los años de medición, diferenciando los distintos niveles económicos según figuras geométricas expresadas en leyenda del margen de la figura.....	171
FIGURA 19. Serie temporal en 15 años de prevalencia de alergia en mujeres representando el porcentaje de prevalencia en ordenadas y en abcisas los años de medición, diferenciando los distintos niveles económicos según figuras geométricas expresadas en leyenda del margen de la figura.....	171
FIGURA 20. Biplot de representación de la distribución de la prevalencia de alergia por niveles económicos, situándose cada uno de los puntos de la muestra ordenados por edad, sexo y niveles según la tabla 12, siendo los ejes los años en que se toman las muestras.....	172
FIGURA 21. Serie temporal de evolución de la prevalencia de alergia según ocupación laboral en 25 años, representando en ordenadas el porcentaje de prevalencia y en abcisas cada uno de los años de medición y diferenciando los distintos grupos de ocupación según colores de la leyenda.....	175
FIGURA 22. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	176
FIGURA 23. Promedio de evolución de prevalencia de alergia en 25 años, representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos tipos de ocupación diferenciándose por cifras totales y sexo.....	177
FIGURA 24. Biplot representado la evolución de la prevalencia de alergia en los años indicados en ejes según tipo de ocupación.....	178

FIGURA 25. Evolución de la prevalencia de alergia según jornada laboral, representando en abcisas cada uno de los tipos de jornada subdivida por sexos y con distintos colores el año de medición. En ordenadas se muestra la cifra de prevalencia.....	180
FIGURA 26. Diagrama de barras con el porcentaje de prevalencia de alergia en ordenadas y en abcisas los diferentes tipos de jornada con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	181
FIGURA 27. Biplot de representación de prevalencia de alergia por tipo de jornada.....	182
FIGURA 28. Promedio de la prevalencia del asma en 12 años por grupos de edad incluidos en divisiones de cifras totales, varones y mujeres.....	186
FIGURA 29. Desglose por años de realización de encuesta (desglosados según color de la leyenda en 4 puntos: 2006, 2011, 2014, 2017) de la evolución de la prevalencia de asma, con el porcentaje de prevalencia representado en ordenadas y en abcisas cada uno de los grupos de edad y sexo (as – ambos sexos-, v -varones- y m –mujeres-).....	187
FIGURA 30. Dibujo en forma de relieve de la prevalencia de asma. Porcentaje de prevalencia representado en ordenadas y en abcisas cada uno de los grupos de edad y sexo (as –ambos sexos -, v -varones- y m –mujeres-). La gráfica dibuja una especie de “V” entre los grupos más jóvenes y los grupos de mayor edad.....	188
FIGURA 31 Diagrama de barras con el porcentaje de prevalencia de asma en ordenadas y en abcisas los diferentes grupos de edad con dos divisiones: varones y mujeres; y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	189
FIGURA 32. Biplot de la evolución de la prevalencia del asma. Representa sobre los ejes a las variables que son cada uno de los grupos de edad y sexo y los ejes son el tiempo en que se efectúan las observaciones con la divisiones representado las cifras de prevalencia.....	190
FIGURA 33. Promedio de la prevalencia del asma en 12 años (desglosados según color de la leyenda en 4 puntos (2006, 2011, 2014, 2017) por tipos de población incluidos en divisiones de cifras totales, varones y mujeres. El porcentaje de prevalencia representado en ordenadas.....	195
FIGURA 34. Diagrama de barras con el promedio del porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de población con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	196
FIGURA 35. Promedio de la prevalencia del asma en 12 años por tipo de población incluidos en divisiones de cifras totales, varones y mujeres.....	197
FIGURA 36. Biplot de representación de evolución de asma según tipo de población.....	198
FIGURA 37. Evolución de la prevalencia de asma en 12 años, desglosados según color de la leyenda en 4 puntos (2006, 2011, 2014, 2017) de realización de encuesta, con el porcentaje de prevalencia representado en ordenadas y en abcisas cada uno de los grupos de nivel de estudios y sexo.....	200
FIGURA 38. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de niveles de estudios con 3 divisiones: cifras totales, varones y mujeres, y a su vez	

cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	201
FIGURA 39. Promedio de la prevalencia del asma en 12 años por nivel de estudios incluidos en divisiones de cifras totales, varones y mujeres.....	202
FIGURA 40. Biplot de representación de evolución de asma según nivel de estudios.....	203
FIGURA 41. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes niveles económicos con 3 divisiones: ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	205
FIGURA 42. Promedio de la prevalencia del asma en 12 años según nivel económico.....	207
FIGURA 43 Biplot de representación de evolución de asma según nivel económico.....	209
FIGURA 44. Promedio de la prevalencia del asma en 12 años por tipos de ocupación incluidos en divisiones de cifras totales, varones y mujeres.....	211
FIGURA 45. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realiza la encuesta (relación de año con su color en la leyenda del pie de la figura).....	212
FIGURA 46. Biplot de representación de evolución de asma según ocupación.....	213
FIGURA 47. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).....	215
FIGURA 48. Promedio de la prevalencia del asma con respecto a jornada laboral incluyendo divisiones de cifras totales, varones y mujeres.....	216
FIGURA 49. Biplot de representación de evolución de asma según tipo de jornada laboral.....	218
FIGURA 50. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia en 25 años por edad y sexo.....	220
FIGURA 51. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia según tipo de población junto por edad y sexo.....	223
FIGURA 52. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia en niveles de estudios y por edad y sexo.....	224
FIGURA 53. Representación gráfica de los conglomerados en que queda dividida la distribución del asma por edad y sexo.....	227
FIGURA 54. Representación gráfica de los clusters en que queda dividida la distribución del asma según tipo de población junto por edad y sexo.....	228
FIGURA 55. Representación gráfica de los conglomerados en que queda dividida la distribución del asma por nivel de estudios junto por edad y sexo.....	230
FIGURA 56. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave de la tabla 45.....	235

FIGURA 57. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.....	237
FIGURA 58. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.....	239
FIGURA 59. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.....	242
FIGURA 60. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave de la tabla 49.....	243
FIGURA 61. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.....	247
FIGURA 62. Codificación en mapa de comunidades autónomas.....	249
FIGURA 63. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.....	253
FIGURA 64. Gráfica de sedimentación de componentes principales de actividad física con alergia y asma. En ordenadas se expresa la varianza y en abcisas cada uno de los componentes. El asma y la alergia en 2011 se explican dentro del grupo mayoritario que corresponde al componente principal 1.....	254
FIGURA 65. Codificación en mapa de comunidades autónomas.....	256
FIGURA 66. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color verde se muestra el conglomerado 1 y en amarillo el conglomerado 2.....	257
FIGURA 67. Dendrograma de distribución de conglomerados de comunidades autónomas en 2011 respecto a factores de salud, alergia, asma.....	258
FIGURA 68. Gráfica de influencias biplot bidimensional de variables y clusters sobre componentes principales.....	259
FIGURA 69. Gráfica de sedimentación de componentes principales de actividad física con alergia y asma en 2017.....	265
FIGURA 70. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.....	265
FIGURA 71. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color azul oscuro se muestra el conglomerado 1 y en color azul claro el conglomerado 2.....	268
FIGURA 72. Dendrograma de la distribución de comunidades autónomas en relación con factores de salud, alergia y asma.....	268
FIGURA 73 Gráfica de influencias biplot bidimensional de variables y conglomerados sobre los componentes principales.....	269
FIGURA 74. Codificación en mapa de comunidades autónomas.....	272
FIGURA 75. Gráfica de sedimentación de componentes principales.....	276
FIGURA 76. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.....	276
FIGURA 77. Círculo de correlaciones de las variables factores de salud, alergia y asma.....	277

FIGURA 78. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color ocre se muestra el conglomerado 1 y en color naranja el conglomerado 2.....	279
FIGURA 79. Gráfica de influencias biplot bidimensional de variables (factores de salud, nominados con X y numerados) y observaciones (conglomerados numerados con el 1 ó el 2) sobre componentes principales.....	280
FIGURA 80. Dendrograma de jerarquización de comunidades autónomas en relación con factores de salud.....	282
FIGURA 81. Gráfica de puntuaciones con algoritmo MASS. Clasificación por conglomerados de comunidades autónomas en relación con factores de salud.....	283
FIGURA 82. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados según los colores de la leyenda.....	284
FIGURA 83. Varianza de estadística descriptiva de factores de salud numerada según clave de comunidades autónomas.....	285
FIGURA 84. La figura muestra la distribución de la prevalencia de alergia y asma a nivel nacional mediante diagrama de caja y diagrama de densidad. En el caso de alergia el diagrama de densidad muestra claramente la normalidad de la distribución y el diagrama de cajas corrobora esta distribución normal con un único valor atípico. En el caso del asma la distribución normal también se confirma con una ligera irregularidad.....	287
FIGURA 85. Diagrama de líneas mostrando en abcisas las comunidades autónomas numeradas de 1 a 20 y en ordenadas la prevalencia de alergia (X17) y de asma (X18).....	288
FIGURA 86. Codificación en mapa de comunidades autónomas.....	289
FIGURA 87. Diagrama de barras mostrando en abcisas las comunidades autónomas numeradas de 1 a 20, junto con la prevalencia de alergia y asma de cada una de ellas en 2011.....	289
FIGURA 88. Diagrama de barras mostrando en abcisas las comunidades autónomas numeradas de 1 a 20, junto con la prevalencia de alergia y asma de cada una de ellas en 2017.....	290
FIGURA 89. Mapa de prevalencia de alergia por comunidades autónomas según color de la leyenda.....	290
FIGURA 90. Mapa de evolución de prevalencia de alergia en España según los colores de la leyenda.....	292
FIGURA 91. Mapa de prevalencia de asma en comunidades autónomas según color de la leyenda.....	293
FIGURA 92. Mapa de evolución de prevalencia de asma en España según los colores de la leyenda.....	294
FIGURA 93 Diagrama de barras mostrando las prevalencias de asma y alergia por comunidades autónomas en los años 2011 y 2017.....	295
FIGURA 94. Evolución de la prevalencia de alergia y asma por comunidades autónomas en los años 2011 y 2017.....	295
FIGURA 95. Codificación en mapa de comunidades autónomas.....	296
FIGURA 96. Gráfico de sectores mostrando la proporción de alergia según tamaño de los cuadros por comunidades autónomas en los años 2011 y 2017.....	297

FIGURA 97. Gráfico de sectores mostrando la proporción de asma según tamaño de los cuadros por comunidades autónomas en los años 2011 y 2017298

LISTA DE TABLAS

TABLA 1. Clasificación de la evolución de la prevalencia de alergia en España desde 1993 a 2017, tanto por cifras totales de ambos sexos, como por separado en varones y mujeres añadiendo los grupos de edad y mostrando los datos de cada uno de los años de medición.....	141
TABLA 2. Cifras de estadística descriptiva de prevalencia en 25 años por grupos de edad y sexo.....	148
TABLA 3. Clave numérica para identificar cada una de las variables que aparecen en la figura 6 y que corresponden a diferentes nichos de edad y sexo.....	150
TABLA 4. Evolución de la prevalencia de alergia desde 1993 a 2017 utilizando como variables los tipos de población según números de habitantes y ofreciendo 3 divisiones: cifras totales, varones y mujeres.....	152
TABLA 5. Estadística descriptiva de la evolución de la alergia en 25 años con respecto al tipo de población en el que vive el alérgico.....	154
TABLA 6. Clave numérica para identificar cada una de las variables que aparecen en la figura 11 y que corresponden a diferentes nichos de edad y sexo.....	161
TABLA 7. Prevalencia de alergia en 12 años diferenciando grupos de niveles de estudios junto con sexo.....	162
TABLA 8. Estadística descriptiva de la evolución de la prevalencia de alergia en 12 años con respecto a nivel de estudios y sexo.....	165
TABLA 9. Clave para entender la figura 14 con las variables numeradas en razón de sexo y nivel de estudios.....	167
TABLA 10. Evolución en 17 años de la prevalencia de alergia según capacidad económica.....	168
TABLA 11. Estadística descriptiva de la evolución de la prevalencia de alergia en 17 años con respecto al nivel económico y sexo.....	170
TABLA 12. Clave numérica para identificar cada una de las variables que aparecen en la figura 20 y que corresponden a diferentes nichos de edad y sexo junto con nivel económico.....	175
TABLA 13. Evolución en 25 años de la prevalencia de alergia según ocupación laboral.....	176
TABLA 14. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años con respecto a situación laboral y sexo.....	179
TABLA 15. Clave numérica con todos los nichos de edad, sexo y ocupación.....	180
TABLA 16. Prevalencia de alergia en 12 años según jornada laboral.....	182
TABLA 17. Estadística descriptiva de la evolución de la prevalencia de alergia en 12 años con respecto a tipo de jornada laboral.....	183
TABLA 18. Clave de alergia y tipo de jornada laboral con todos los nichos de edad, sexo y ocupación.....	184
TABLA 19. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (15-24 años, 25-34 años, 35-44 años, 45-54 años, 55-64 años, 65-74 años, más de 75 años).....	186
TABLA 20. Valores de estadística descriptiva a lo largo 12 años (2006, 2011, 2014, 2017) sobre la prevalencia de asma, distribuidos por divisiones por totales de ambos sexos, varones y mujeres	

desglosados en grupos de edad (15-24 años, 25-34 años, 35-44 años, 45-54 años, 55-64 años, más de 75 años).....	187
TABLA 21. Clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.....	192
TABLA 22. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (cifras totales, 15-24 años, 25-44 años).....	195
TABLA 23. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (45-64 años, más de 65 años).....	196
TABLA 24. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 por tipos de población.....	197
TABLA 25. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a tipo de población y sexo.....	199
TABLA 26. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y los tipos de población y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.....	200
TABLA 27. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 por nivel de estudios.....	201
TABLA 28. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a nivel de estudios y sexo.....	203
TABLA 29. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y los niveles de estudios y numerando cada una de estas divisiones.....	205
TABLA 30. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 según nivel económico.....	207
TABLA 31. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto al nivel económico incluyendo divisiones de cifras totales, varones y mujeres	208
TABLA 32. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el nivel económico y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.....	210
TABLA 33. Cifras de prevalencia de asma por ocupación, en los cuatro años de medición, y desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (15-24 años, 25-44 años, 45-64 años, más de 65 años).....	212
TABLA 34. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a ocupación y sexo.....	213
TABLA 35. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el tipo de ocupación y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.....	215
TABLA 36. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 según tipo de jornada laboral.....	216

TABLA 37. Estadística descriptiva de la evolución de la prevalencia de asma con respecto a jornada laboral incluyendo divisiones de cifras totales, varones y mujeres.....	218
TABLA 38. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el tipo de jornada laboral y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.....	219
TABLA 39. Clave numérica que muestra ordenados los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.....	221
TABLA 40. Clave que muestra ordenados numéricamente los diferentes nichos de población, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.....	224
TABLA 41. Clave que muestra ordenados numéricamente los diferentes nichos de niveles de estudios, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.....	227
TABLA 42. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico, con sus respectivas numeraciones de pertenencia a un determinado cluster (localizadas en la zona de verde más intenso). Así mismo aparecen las cifras de prevalencia.....	228
TABLA 43 Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado. Así mismo, aparecen las cifras de prevalencia.....	231
TABLA 44. Clave que muestra ordenados numéricamente los diferentes nichos de niveles de estudios, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de verde más intenso). Así mismo, aparecen las cifras de prevalencia.....	234
TABLA 45. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico.....	236
TABLA 46. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con tipo de población que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.....	238
TABLA 47. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con nivel de estudios que son utilizados para el análisis estadístico. Incluye cifras de prevalencia...	242
TABLA 48. Muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico e incorpora cifras de prevalencia.....	243
TABLA 49. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el tipo de población que son utilizados para el análisis estadístico e incorpora cifras de prevalencia.....	249

TABLA 50. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el nivel de estudios que son utilizados para el análisis estadístico e incluye la prevalencia.....	248
TABLA 51. Clave para identificar las variables utilizadas de actividad física.....	250
TABLA 52. Clave de codificación para identificar a las comunidades autónomas.....	251
TABLA 53. Estadística descriptiva de 2011 de las variables en relación con todas las comunidades autónomas. Estadística descriptiva de la prevalencia de las variables con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (desviación típica, coeficiente de variación).....	252
TABLA 54 Resumen de la matriz de correlaciones con respecto a la alergia y el asma en 2011.....	253
TABLA 55. Resumen de los componentes principales ligados a la alergia y el asma en 2011. Proporción de varianza explicada por los principales componentes.....	255
TABLA 56. Codificación de las comunidades autónomas con su pertenencia al conglomerado concreto.....	258
TABLA 57. Estadística descriptiva de 2017 de la prevalencia de las variables en relación con todas las comunidades autónomas. Estadística descriptiva con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (desviación típica, coeficiente de variación).....	263
TABLA 58. Resumen de la matriz de correlaciones de factores de salud y alergia y asma en 2017....	264
TABLA 59. Comparativa (coeficientes de correlación de Pearson) de factores de actividad física, alergia y asma.....	265
TABLA 60. Análisis de componentes principales de 2017 de las variables en todas las comunidades autónomas.....	266
TABLA 61. Resumen de los componentes principales ligados a la alergia y el asma. Proporción de varianza explicada por los principales componentes.....	268
TABLA 62. Análisis de conglomerados de 2017 de las variables en todas las comunidades autónomas.....	269
TABLA 63. Codificación de factores de salud con su numeración.....	273
TABLA 64. Referencia de las comunidades autónomas con su numeración.....	273
TABLA 65. Estadística descriptiva con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (coeficiente de variación, desviación típica).....	274
TABLA 66. Resumen de la matriz de correlación de 2017 de las variables en todas las comunidades autónomas.....	276
TABLA 67. Tabla de análisis de componentes principales de factores de salud (peso, hábito tabáquico, actividad física en trabajo, sedentarismo en ocio, tipo de desayuno).....	277
TABLA 68. Tabla de codificación de comunidades autónomas con sus conglomerados correspondientes.....	281
TABLA 69. Tabla de codificación de comunidades autónomas con sus conglomerados correspondientes.....	286

I. INTRODUCCIÓN

1.1. Resumen histórico de la alergia

Las enfermedades alérgicas se encuentran entre las causas más frecuentes de patología crónica, su prevalencia está aumentando desde hace años (1, 2) y suponen una carga adicional y cada vez en mayor grado a los costes sanitarios (3).

Hay que explicar que el término “alergia” fue introducido en 1906 por Von Pirquet, al mostrar que en las reacciones de hipersensibilidad, el antígeno inducía un cambio en la reactividad habitual. Otro de los términos utilizados para explicar el fenómeno alérgico es la palabra “atopia”, introducido por Coca y Cooke, la cual deriva del griego *atopos* que significa “fuera de lugar” y suele emplearse para describir a pacientes que tienen una predisposición hereditaria a producir anticuerpos de tipo IgE frente a los alérgenos ambientales comunes. La atopia sería la tendencia para producir inmunoglobulina E específica (IgE) tras la exposición a un alérgeno, incluyendo enfermedades como el asma, la rinitis alérgica, la alergia alimentaria y la dermatitis atópica (4). Con el tiempo, la palabra alergia ha pasado a emplearse como sinónimo de enfermedad mediada por IgE con manifestaciones que se extienden por toda la economía corporal (vías aéreas, pulmón, piel, aparato digestivo, etc). Las enfermedades alérgicas incluyen a la rinitis, la rinoconjuntivitis, el asma, la alergia alimentaria, la alergia a los venenos de insectos, la alergia a fármacos y las enfermedades cutáneas alérgicas (urticaria, eccema, etc). Algunas enfermedades alérgicas se manifiestan a través de mecanismos independientes de la IgE como sería el caso de las dermatitis de contacto o de las neumonitis por hipersensibilidad.

Cuando los alérgenos se introducen en el cuerpo de una persona predispuesta a la alergia, el organismo genera anticuerpos específicos (inmunoglobulinas IgE) para combatir al que consideran un agente invasor. Si continúa la exposición al alérgeno, estos anticuerpos específicos interactúan con unas células denominadas mastocitos y éstas liberan sustancias como la histamina, que son las que producen los síntomas propios de la alergia. El desarrollo de alergia que implica la formación de anticuerpos de inmunoglobulina E depende tanto de factores genéticos complejos como de influencias ambientales. El proceso por el cual una persona se vuelve

alérgica a una sustancia empieza con la sensibilización. Se inicia con la entrada del antígeno su captación y procesamiento por las células presentadoras de antígeno (célula dendrítica, macrófagos) y presentación a los linfocitos que serán capaces de producir una respuesta inmune específica, por medio de los linfocitos B y dar lugar a anticuerpos IgE contra una sustancia inhalada, ingerida o inyectada (5). Posteriormente esta IgE se adhiere a los basófilos circulantes en la sangre o a los mastocitos situados en la piel y mucosas del tracto gastrointestinal y del sistema respiratorio. Cuando se produce una nueva exposición al alérgeno, la unión de éste a la IgE fijada en la superficie de mastocitos y basófilos producirá la apertura de dichas células con la consiguiente liberación de mediadores responsables de las reacciones alérgicas. Casi todos los órganos pueden estar afectados y las manifestaciones comprenden: conjuntivitis (ojos), rinitis (nariz), asma (pulmones), urticarial/dermatitis (piel), etc.

Se ha visto que existe un papel de la genética en la tendencia a producir IgE, favoreciendo la presencia de una inflamación crónica en los territorios afectos, por ejemplo a nivel cutáneo o bronquial. En esta inflamación crónica están implicados mastocitos, basófilos, linfocitos T específicos del alérgeno y eosinófilos activados. Un gran número de los genes que codifican las citocinas y los receptores que regulan la inflamación alérgica se encuentran agrupados en un segmento corto del cromosoma 5q y en el 11q13, y en algunas familias se ha visto una relación entre las concentraciones séricas totales de IgE y estos genes (6). Por ejemplo, los polimorfismos en el gen que codifica la linfopoyetina estromal tímica (TSLP) y su receptor influyen en el riesgo de dermatitis atópica, alergia alimentaria y el asma, mientras que los polimorfismos en los genes que codifican la interleucina 33 (IL-33) y su receptor se asocian con un mayor riesgo de dermatitis atópica y asma (7). Estos estudios apoyan la existencia de un conjunto compartido de polimorfismos que predisponen a los pacientes a la dermatitis atópica y promueven el desarrollo posterior de otras afecciones alérgicas (8).

Por otro lado, la producción de IgE específica de muchos de los alérgenos del polen y animales domésticos se ha relacionado con ciertas moléculas MHC de clase II. Estos datos indican que la herencia en las enfermedades alérgicas es multifactorial. Hay

otros factores que influyen en que una persona nacida en una familia atópica desarrolle estas enfermedades (9). De hecho, hay autores que refieren que en las familias en que uno de los padres presenta una dermatitis atópica es más probable que los hijos también la presenten, así como en las familias en las que uno de los padres presenta asma o rinitis alérgica. En contacto con aeroalérgenos, la población no atópica pone en marcha una respuesta de baja intensidad y producen IgG1 e IgG4 específicas del alérgeno, y sus linfocitos T responden con un grado moderado de producción de interferón gamma que es típico de su población celular Th1. Por el contrario, las personas atópicas ponen en marcha una respuesta IgE específica exagerada frente al alérgeno y sus linfocitos T producen citoquinas de tipo Th2 como las interleucinas 4 y 5 (10). Los linfocitos T de todos los recién nacidos son estimulados frente a los alérgenos ambientales comunes y producen una respuesta inmune dominada por Th2 (11). Durante el desarrollo inmune de una persona no atópica habría una desviación a favor de la respuesta Th1 frente a los alérgenos inhalados (desviación inmune). Sin embargo, en los atópicos habría un refuerzo posterior de la inmunidad polarizada en Th2 (12). Probablemente, el principal estímulo para el desarrollo de la inmunidad protectora de tipo Th1 es la exposición microbiana (13).

La marcha atópica es un intento de explicar la alergia a través de la descripción de la historia natural de las enfermedades alérgicas (14). El inicio probablemente resulta de la interrupción de la barrera cutánea debido a defectos intrínsecos de las células epiteliales en un individuo con una predisposición genética y/o ambiental para la inflamación regulada por linfocitos Th2. La exposición a alérgenos a través de la piel inflamada es la ruta principal por la cual los individuos inician la marcha atópica (15). Esta hipótesis está respaldada por datos de modelos animales que indican que la exposición transcutánea a alérgenos promueve el desarrollo de respuestas específicas de células T y B y la posterior enfermedad alérgica. Estos modelos pueden tomar la forma de alteraciones genéticas (como es el caso de los ratones deficientes en filagrina) o irritación de la piel por medios mecánicos o químicos. Una vez que un alérgeno ha entrado en la piel, tiene la oportunidad de interactuar con el sistema inmunológico.

La piel se divide en dos compartimentos inmunes: la epidermis, que contiene predominantemente células de Langerhans y linfocitos T CD8 citotóxicos, y la dermis, que contiene células dendríticas dérmicas y plasmocitoides, macrófagos, mastocitos y subconjuntos de linfocitos innatos y adaptativos. Las células dendríticas (CD) activadas procesan alérgenos, regulan al alza el complejo mayor de histocompatibilidad (MHC) y circulan a los ganglios linfáticos drenantes, donde pueden interactuar con las células T y B *naïve* para promover el desarrollo de respuestas de células T y B específicas de alérgenos (16). Los linfocitos T específicos para alérgenos Th2 regresan a la piel a través de la expresión del receptor de quimiocinas C-C tipo 4 (CCR4), antígeno linfocitario cutáneo (CLA) y otras moléculas. Los linfocitos Th2 también entran en la circulación sistémica, donde pueden ejercer una respuesta efectora en sitios de tejido distantes. Los linfocitos B de memoria recirculan en la sangre y la linfa, mientras que los plasmablastos se albergan en la médula ósea, donde se diferencian en células plasmáticas y producen IgE específica (17).

La inflamación observada en la atopia se asocia con una mayor producción de IL-4, IL-25, IL-33 y linfopoyetina estromal tímica que reclutan células linfoides innatas tipo 2 productoras de IL-5 e IL-13 y contribuyen al desarrollo de la inflamación tipo 2 (17). Las citocinas derivadas del epitelio, como la interleucinas IL-25, IL-33 y la linfopoyetina del estroma tímico (TSLP) reclutan y activan tipos de células innatas, incluidas las células linfoides innatas (ILC) y los basófilos, que producen citoquinas que promueven la activación de las células dendríticas (CD) (18).

Aunque las causas genéticas son determinantes importantes de la alergia, la exposición ambiental contribuye indudablemente a la variabilidad en la expresión de las enfermedades alérgicas en los individuos atópicos (19). El actual paradigma de la patogenia alérgica es que la alergia se desarrolla en individuos con predisposición genética sólo después de que hayan sido expuestos a alérgenos (20). En general, se cree que los aumentos epidémicos en la prevalencia y la gravedad de las afecciones alérgicas observadas en las últimas décadas han ocurrido demasiado rápido para atribuirse sólo a la deriva genética. Como resultado, se ha dedicado un esfuerzo considerable a comprender el papel de la dieta, la higiene, las infecciones,

los alérgenos, la contaminación del aire y otros factores ambientales en la susceptibilidad a las enfermedades alérgicas (21).

La exposición a contaminantes ambientales tanto en lugares abiertos como cerrados, o el desarrollo de infecciones respiratorias leves de origen viral en la primera infancia, pueden aumentar la probabilidad de presentar una sensibilización al alérgeno y al posterior desarrollo de síntomas respiratorios crónicos (22).

El marcado incremento en la prevalencia de las enfermedades alérgicas en países occidentales indica la importancia de las influencias ambientales. Un ejemplo interesante de cambios ambientales en oposición a cambios genéticos es la incidencia de rinitis alérgica estacional y de asma después de la reunificación de Alemania. Estos trastornos eran menos comunes en la Alemania Oriental que en la Alemania Occidental antes de la reunificación (23), mientras que después de la misma la prevalencia de atopia y rinitis alérgica, pero no de asma, se ha incrementado en los niños que pasaron su primera infancia en Alemania Oriental (24,25).

Este fenómeno plantea la posibilidad de que un estilo de vida occidental explique el aumento de la prevalencia y se ha sugerido que en los países occidentales el sistema inmune en desarrollo carece de agentes microbianos que estimulen los linfocitos Th1, debido a que el ambiente está limpio y por el uso extendido de antibióticos (26).

Varios estudios epidemiológicos apoyan esta teoría. Algunos de los datos recientes más convincentes que respaldan el papel de la **exposición microbiana** en el riesgo atópico provienen de estudios que compararon poblaciones pediátricas geográficamente distintas pero genéticamente relacionadas (27). En un estudio de niños en Finlandia y Rusia, una reducción significativa dependiente de la dosis en el riesgo de atopia se asoció con la exposición microbiana y la prevalencia de enterovirus. Por ejemplo, la evidencia de que la colonización bacteriana del tracto gastrointestinal evita la sensibilización atópica surgió de estudios en niños que vivían en países con una prevalencia baja (Estonia) o alta (Suecia) de atopia. Lactobacilli y Eubacteria predominaban en los niños estonios, mientras que Clostridia era más frecuente en niños suecos (28).

Estos hallazgos pueden explicar por qué un estilo de vida agrícola, en el que existe una mayor exposición a bacterias en los establos donde se guarda el ganado, es protector frente al desarrollo de la enfermedad atópica (29). En otro estudio de niños amish y huteritas residentes en granjas, la prevalencia de asma y sensibilización alérgica fue cuatro y seis veces menor en los amish (que siguen prácticas agrícolas más tradicionales). Este hallazgo se correlacionó con el polvo en los hogares amish que tenían 6,8 veces más endotoxina (30).

Otros factores que pueden influir son la posición ocupada en el número de hermanos, puesto que las enfermedades atópicas son menos frecuentes en los hermanos pequeños cuando los mayores han pasado enfermedades infecciosas como sarampión y hepatitis A, lo que indica que la estimulación inmune repetida por virus o bacterias puede ser protectora (31). Uno de los beneficios de los programas modernos de vacunación es la reducción de las infecciones graves durante la infancia; sin embargo, se ha sugerido que este cambio puede ser el responsable de los recientes incrementos de la alergia y el asma. Aunque los mecanismos que producen estos efectos son desconocidos, algunos investigadores han referido la hipótesis de que las infecciones respiratorias frecuentes estimulan la producción de interferón gamma en la infancia, y de la presencia de esta citocina durante el estado crítico del desarrollo del sistema inmunológico (32).

El desarrollo de enfermedades alérgicas puede estar relacionado con alteraciones en el órgano diana. Por ejemplo, los cofactores necesarios para la aparición de una crisis de asma pueden incluir las infecciones por virus respiratorios y la exposición a contaminantes atmosféricos (33). El papel de los **factores ambientales** en relación con el asma y la alergia ha ido cobrando interés creciente (34). Ha habido una preocupación general respecto a que los cambios en la contaminación ambiental hayan sido la base del incremento del asma (35). La cuestión más importante es si la exposición a largo plazo a la contaminación atmosférica influye en la prevalencia del asma. Esto ha sido evaluado en estudios por áreas en los que se ha comparado la prevalencia de asma en toda una ciudad o en ciudades con mezclas diferentes de contaminación (36).

Es probable que los cambios en los ambientes cerrados hayan tenido una gran influencia en el incremento de la prevalencia de enfermedades alérgicas, en su mayor parte a causa de que la población de las sociedades industrializadas pasa la mayor parte de su tiempo en lugares cerrados. Se ha prestado menos atención al ambiente interior, en el cual las personas pasan la mayor parte de su tiempo (37). Los factores ambientales que pueden provocar crisis de broncoespasmo son los gases irritantes, partículas inorgánicas, alérgenos e infecciones (38). El aumento de la humedad en las casas con aire acondicionado fomenta, por ejemplo, el crecimiento de mohos y ácaros del polvo, con lo que se produce un incremento de la exposición al alérgeno (39,40). Los ácaros son artrópodos (clase arácnida) que miden de 0,2 a 0,5 milímetros. Se encuentran habitualmente en espacios cerrados donde hay polvo, generalmente lugares oscuros alejados de la luz solar. Algunos predominan en las casas y otros se encuentran preferentemente en almacenes, granjas o graneros (41). Las especies que más a menudo causan trastornos alérgicos asociados al polvo de las casas son el *Dermatophagoides pteronyssinus* y el *Dermatophagoides farinae*, que se alimentan de las escamas microscópicas que nuestra piel desprende poco a poco en su proceso continuo de renovación (42). Por tanto, suelen encontrarse en los lugares donde abundan los restos de escamas de piel de personas: colchones, almohadas, mantas, edredones, sofás, alfombras, muebles tapizados con tela, peluches, ropa. También existen los conocidos como ácaros de almacenamiento – *Acarus siro*, *Lepidoglyphus destructor* y *Tyrophagus putrescentiae*, entre otros-. Estos se alimentan principalmente de granos y de otras partículas de comida y se hallan en los espacios donde se almacenan alimentos vegetales, especialmente cereales y harinas, y también, en el caso del *Tyrophagus putrescentiae*, en los alimentos ricos en proteínas como el jamón o el queso. En el caso de la alergia a los ácaros, la hipersensibilidad es a las proteínas presentes en las heces de los ácaros o a las partículas en que se desintegra su cuerpo cuando mueren. Al penetrar en las vías respiratorias, la reacción del organismo para protegerse puede causar trastornos como la rinoconjuntivitis y el asma alérgica. Esta exposición puede producirse de manera repentina, cuando una persona, por ejemplo, respira una gran cantidad de partículas procedentes de los ácaros limpiando el polvo o sacudiendo la ropa de cama. En estas circunstancias, puede padecer un ataque agudo de asma o

rinoconjuntivitis. Pero, normalmente, la exposición a este alérgeno se produce de manera paulatina, en pequeñas cantidades pero de manera constante, dado que las personas permanecen en su domicilio durante muchas horas al día (43).

La frecuencia de las crisis de alergia y asma varía de un día para otro como indican las cifras de ingresos hospitalarios o de visitas a urgencias (44). Existe una variación estacional predecible (45) y además hay otras variaciones diarias que se han relacionado con la importancia de la contaminación atmosférica exterior (46,47). Los contaminantes químicos formados principalmente por materia inanimada, se presentan en el aire como gases, vapores, humos, aerosoles (partículas sólidas o pequeñas gotas de líquidos que pueden permanecer en el aire un tiempo determinado) y nieblas (son suspensiones de finas gotas líquidas que se producen por condensación de vapores o por la dispersión en un estado líquido de fluidos que se expanden hasta ocupar el recinto que los contiene) (48,49). Los componentes más abundantes de la contaminación del aire en las zonas urbanas son el dióxido de nitrógeno, el ozono y el material particulado (PM) (50,51). El dióxido de azufre es abundante en las áreas industriales (52). El material particulado contaminante del aire incluye una mezcla de pequeñas partículas de sólidos, líquidos o mezcla de sólidos y líquidos, orgánicos e inorgánicos, naturales y antropogénicos, de composición variable, suspendidos en el aire (53).

Se ha intentado relacionar el efecto de las partículas existentes en atmósferas contaminadas con la aparición de alergia y asma (54, 55). Los contaminantes que han sido determinados con mayor frecuencia en las comparaciones geográficas para relacionar la prevalencia del asma y la contaminación atmosférica son el dióxido de azufre y las partículas aéreas (56, 57). Otros estudios han evaluado los posibles peligros debidos a contaminantes como óxidos de nitrógeno (58) y ozono (59). En general, los valores de óxido de nitrógeno se correlacionan con las concentraciones de partículas de dióxido de azufre y resulta difícil diferenciar sus efectos independientes (60).

En la reunificación de Alemania, un hallazgo relevante fue el de una menor prevalencia de respuestas cutáneas positivas en las zonas contaminadas de las

ciudades de Alemania del Este (61). Parecía haber un exceso de síntomas de irritación (incluyendo bronquitis) en las áreas más intensamente contaminadas de la Alemania del Este equilibrado con una menor prevalencia de asma alérgica (62).

También ha habido una preocupación respecto a que la exposición al tráfico denso podría aumentar el riesgo de asma (63).

Los contaminantes del aire relacionados con el tráfico producidos por los vehículos de motor se componen de una mezcla compleja de partículas (PM, PM10 o PM2.5) y emisiones gaseosas que incluyen óxidos de nitrógeno, monóxido de carbono, ozono y otros contaminantes del aire. Se han asociado con función pulmonar disminuida y desarrollo y exacerbación del asma. Los escapes de automóviles, camiones y plantas de energía son las fuentes más importantes de dióxido de nitrógeno al aire libre, el cual es un precursor de la polución fotoquímica que se encuentra en las regiones urbanas e industriales (64). La exposición al dióxido de nitrógeno se asocia con un aumento de las visitas a urgencias, las sibilancias y el uso de medicamentos entre los niños con asma. Varios estudios poblacionales sugerían relación entre el asma y contaminantes producidos por el tráfico (65, 66) y que había una mayor prevalencia de asma y alergia en niños que vivían en calles con mucho tráfico (67).

Sin embargo, otros trabajos admiten que el aumento repentino de los niveles de contaminantes ambientales, debido al desarrollo industrial y el tráfico de vehículos de motor urbano, ha afectado la calidad del aire, pero no encuentran una relación tan directa con enfermedades alérgicas (68), o no detectan una relación consistente (69).

Ya que se admite que los contaminantes ambientales pueden actuar sobre el sistema inmunológico y respiratorio, se trataría de encontrar los mecanismos inmunológicos que explicasen que la contaminación del aire o del agua pueda potenciar de alguna manera tanto las exacerbaciones de asma como la sensibilización a los alérgenos. Así, el dióxido de nitrógeno al inducir la nitración de alérgenos favorecería las respuestas linfocíticas de Th2, lo que conduciría a una mayor alergenicidad del polen en áreas contaminadas y un mayor riesgo para la salud humana (70). Además, el estrés oxidativo en pacientes con enfermedades respiratorias puede inducir inflamación eosinofílica en las vías respiratorias,

aumentar la sensibilización alérgica atópica y aumentar la susceptibilidad a infecciones (71).

Así mismo, los factores atmosféricos como las tormentas pueden desencadenar crisis de asma (72). La polinosis o alergia al polen es una enfermedad causada por una reacción alérgica frente a los distintos tipos de polen que, presentes en la atmósfera, penetran en el organismo a través de las mucosas expuestas al aire y producen procesos respiratorios como la rinitis y el asma. Normalmente, las precipitaciones provocan un descenso de los niveles de polen ya que al humedecer el ambiente, las partículas de polen en el aire pesan más, descienden y se caen por gravedad. Y, al no estar a la altura óptima para ser respiradas, afectan menos a los alérgicos (73). El cambio climático y las alteraciones ambientales están modificando el entorno natural lo que afecta a la polinosis. Sin embargo, el asma por tormenta se caracteriza porque durante estos fenómenos de precipitaciones intensas y virulentas, la lluvia y el viento rompen los granos de polen, lo que puede producir que haya proteínas alérgicas en pequeñas partículas suspendidas (74). Esta situación puede incrementar los síntomas de alergia y asma a pesar de que no haya una gran concentración de polinosis. La lluvia puede aumentar la concentración de hongos y además romper los granos de polen y liberar partículas alérgicas que también producen síntomas a estos pacientes, que podrían ser incluso más intensos que los provocados por los granos completos. Según las estimaciones, durante los primeros 20-30 minutos de una tormenta, las personas que padecen alergia pueden inhalar una alta concentración del material alérgico que se dispersa en la atmósfera, que a su vez puede inducir reacciones asmáticas que pueden ser muy graves. Se cree que el asma por tormentas eléctricas se desencadena por una combinación única de niveles altos de polen y cierto tipo de tormenta (75, 76). Los granos de polen son arrastrados por el viento y transportados a largas distancias. El polen puede romperse y liberar partículas diminutas que se concentran en las ráfagas de viento que llegan justo antes de una tormenta. Estas partículas son lo suficientemente pequeñas como para ser inhaladas profundamente en los pulmones y pueden desencadenar síntomas de asma dificultando la respiración (77, 78).

En resumen, los estudios epidemiológicos indican que en las alteraciones alérgicas se solapan la expresión clínica, el vínculo genético y los mecanismos patogénicos. A pesar de estas similitudes, es probable que tanto los factores genéticos como los ambientales influyan en el riesgo de sensibilización al alérgeno, y afecten a la expresión clínica de la alergia (79, 80, 81).

Las enfermedades alérgicas, y sobre todo el asma, son una de las principales causas de visitas en consultas externas e ingresos hospitalarios, con la consiguiente carga financiera (82, 83, 84). Otros costes atribuibles incluyen las pérdidas por bajas laborales, las limitaciones de la actividad y el estrés psicológico (85). La prevalencia y la morbilidad de las enfermedades alérgicas y del asma no se distribuye de manera uniforme, ya que los países desarrollados presentan, en general, tasas más elevadas (86) que los países del Tercer Mundo. La percepción es que tanto el asma y la dermatitis atópica como las rinitis alérgicas presentan un aumento de las prevalencias. Estas tendencias, además del gran incremento de la alergia observado en los países desarrollados en comparación con los del Tercer Mundo, plantean la posibilidad de que los factores causantes se asocien con los cambios en los estilos de vida impulsados por la sociedad industrial moderna.

En general, el asma se inicia en la más temprana infancia; la mayoría de los niños presentan su primer episodio de asma antes de los tres años. La prevalencia del asma está relacionada de forma importante con la atopia. Más de la mitad de los niños asmáticos no presentan síntomas en la edad adulta; el desarrollo temprano de asma y la ausencia de rinitis alérgica indican un buen pronóstico. Sin embargo, es probable que los niños con asma grave continúen presentando problemas cuando llegan a la edad adulta (87, 88).

La mayoría de los adultos con asma inician síntomas respiratorios en la infancia, o con menos frecuencia, entre los 20 y los 40 años de edad. La exposición a alérgenos transportados por el aire o a alérgenos químicos en el lugar de trabajo debe ser considerada en la evaluación del asma del adulto. A pesar de los distintos desencadenantes del asma alérgica y no alérgica, la inflamación de la vía aérea y las manifestaciones clínicas son similares. Los pacientes de mediana edad o los ancianos que inician síntomas sugerentes de asma deben ser evaluados para descartar otras enfermedades, incluyendo en el diagnóstico diferencial tanto la

enfermedad pulmonar obstructiva crónica (EPOC), como las obstrucciones mecánicas de la vía aérea, el reflujo gastroesofágico o las dolencias respiratorias de origen cardíaco. En general, el asma que aparece en la edad adulta es menos probable que se resuelva de forma espontánea que el asma que se inicia en la infancia (89).

Existe una gran variabilidad en la prevalencia del asma a nivel mundial y se detecta una estabilización o descenso que ha ocurrido principalmente en los países desarrollados, mientras que en los países en desarrollo continúa aumentando. Las causas de estos cambios aún son desconocidas (90). Una de las explicaciones con más aceptación es que factores ambientales, en especial motivados por la «occidentalización» de la forma de vida, están modificando la respuesta inmunológica y se especula que la etapa prenatal o los primeros años de vida son los periodos más susceptibles para que ocurran estos cambios. En este contexto, una de las teorías más aceptadas, denominada «hipótesis de la higiene» (26), postula que la ausencia de la exposición habitual a microorganismos hace que el sistema inmunológico tenga una inmadurez relativa y que, por lo tanto, sus respuestas no sean de tipo protector (promovidas por la población Th1 de linfocitos T) sino de tipo alérgico (promovidas por la población Th2 de linfocitos T). Otras hipótesis están relacionadas con exposiciones a diversos factores como humo de tabaco, alérgenos, contaminantes atmosféricos, contaminantes intramuros o dieta de tipo occidental. Un mecanismo común a través del cual algunos de estos factores podrían estar operando son las modificaciones epigenéticas como la metilación del ADN o la acetilación de histonas, que promueven silenciamiento o cambios en la expresión de genes, sin implicar cambios en la secuencia del ADN (90).

La rinitis alérgica es la manifestación más común de la atopia, afecta cerca del 20% de la población y se sitúa entre los problemas de salud más importantes (91). La mitad de los pacientes con rinitis tienen conjuntivitis asociada. Así, la rinitis alérgica tiene un impacto significativo en la salud pública y representa un coste económico elevado tanto de prescripción médica como de venta directa de medicamentos sin receta. Entre las complicaciones asociadas con las rinitis alérgicas se incluyen la pérdida de gusto y del olfato, las alteraciones del sueño, la sinusitis, las disfunciones

de la trompa de Eustaquio y la otitis media. Los factores de riesgo para el desarrollo de la rinitis alérgica incluyen el asma, la historia familiar de alergia, los valores elevados de IgE sérica, la exposición al humo del tabaco y a los alérgenos ambientales (92). La máxima prevalencia de rinitis alérgica se sitúa en la adolescencia hasta la tercera década de la vida, y es típico que los síntomas disminuyan al alcanzar una mediana edad. Aproximadamente el 20% de los niños con rinitis alérgica evoluciona favorablemente, aunque pueden persistir otras manifestaciones de atopia como el asma. La asociación clínica entre el asma y la rinitis alérgica está bien establecida, con hasta tres cuartas partes de las personas con asma que informan síntomas de rinitis. Esta asociación se mantiene después de controlar la IgE total, los antecedentes parentales de asma y la sensibilización a los alérgenos, lo que sugiere que la coexistencia de rinitis alérgica y asma no es únicamente atribuible a la predisposición atópica. La rinitis alérgica también se correlaciona positivamente con la gravedad del asma, y el tratamiento de la rinitis alérgica mejora el control del asma. En conjunto, estas observaciones indican que las vías respiratorias superiores e inferiores se comportan como una unidad fisiológica y fisiopatológica y han llevado a recomendaciones integrales del panel de guías de rinitis alérgica y su impacto en el asma, que enfatizan la importancia del tratamiento apropiado de rinitis alérgica en individuos con asma (14).

La dermatitis atópica es, a menudo, la primera manifestación de la alergia en un niño atópico. Es especialmente frecuente en niños y adolescentes. Como en el caso de otras enfermedades alérgicas, la prevalencia de la dermatitis atópica ha aumentado en los últimos años. El pronóstico de la dermatitis atópica es, en general, favorable; los síntomas desaparecen por completo aproximadamente en una tercera parte de los casos que se iniciaron en la infancia. La dermatitis se asocia a menudo con otras enfermedades alérgicas (93,94).

La marcha atópica describe la historia natural de las enfermedades alérgicas. El inicio probablemente resulta de la interrupción de la barrera cutánea debido a defectos intrínsecos de las células epiteliales en un individuo con una predisposición genética y/o ambiental para la inflamación tipo 2. Los defectos en la proteína de barrera epidérmica filagrina y las mutaciones de pérdida de función en SPINK5 y el

gen que codifica la corneodesmosina a menudo se citan como un ejemplo por excelencia de tal defecto de barrera (14).

Las mutaciones de pérdida de función en el gen de la filagrina se correlacionan con la susceptibilidad y la gravedad del asma en pacientes con enfermedad atópica, pero no en aquellos sin ella, lo que indica que la inflamación de la piel es necesaria para la sensibilización alérgica. A pesar de estos datos, no todos los pacientes con enfermedad atópica desarrollan asma, y no todos los pacientes con asma tienen enfermedad atópica previa. Un análisis retrospectivo reciente de dos cohortes de nacimiento encontró ocho patrones separados de progresión de la enfermedad atópica. Aunque limitado por el tamaño de la cohorte y la encuesta de los padres como método de informe, este estudio encontró que el 10,5% de los encuestados siguió el patrón tradicional de la marcha atópica, mientras que el 15,5% tenía dermatitis atópica persistente, el 5,7% tenía sibilancias sin dermatitis atópica y el 9,6% tenía rinitis sin dermatitis atópica. En conjunto, estos hallazgos indican que la marcha atópica no está presente en todos los individuos atópicos, y, en particular, aquellos con enfermedad de inicio en adultos (14).

El desarrollo de la alergia alimentaria puede tener también un impacto significativo en la gravedad de la atopía (95). Entre un tercio y la mitad de los niños remitidos a clínicas de asistencia terciaria de alergia a causa de una dermatitis atópica intensa presentan una alergia alimentaria subyacente (96). Eliminando de la dieta los alimentos responsables puede producirse una rápida mejoría de la afectación cutánea, aunque coexistan la rinitis y el asma. Entre un 6-8% de niños menores de 3 años y el 2-6% de adultos experimenta alergia alimentaria (97). En un estudio prospectivo en recién nacidos se encuentra un 8% de alérgicos a alimentos (98). Las respuestas de IgE específicas de los alimentos se pueden detectar en los primeros meses de vida y alcanzar un máximo de prevalencia de aproximadamente el 10% a 1 año de edad. El hecho de que la sensibilización ocurra antes de la ingestión de alimentos en la mayoría de los casos sugiere que la sensibilización a los alimentos ocurre a través de la exposición a través de la piel inflamada, a diferencia del tracto gastrointestinal (14).

La dermatitis atópica en adultos se ve raramente agravada por la alergia alimentaria, por el contrario las exposiciones ocupacionales contribuyen frecuentemente a la gravedad de la enfermedad. Los individuos atópicos presentan una incidencia muy superior de dermatitis ocupacional, afectando con más frecuencia a las manos. Entre las causas de dermatitis ocupacional se incluyen tanto irritantes como alérgenos. La relación entre la dieta y el asma es un área controvertida que no ha terminado de ser evaluada. Existen pocos trabajos sobre la influencia de la alimentación en el asma (94).

La prevalencia de la anafilaxia está subestimada como consecuencia de la dificultad de recopilar estas graves reacciones alérgicas. Existen muy pocos estudios epidemiológicos sobre la incidencia real de la anafilaxia y no hay unos códigos de registro universalmente (99).

En cuanto a la urticaria, afección cutánea caracterizada por la presencia de habones y que puede ser contemplada como el efecto de la degranulación de los mastocitos cutáneos, los estudios valoran de forma variable su prevalencia (100,101,102,103). El pronunciado aumento global en la prevalencia y gravedad de las enfermedades atópicas durante las últimas décadas es de importancia crítica para la salud de nuestra población. El coste económico de las enfermedades alérgicas es muy superior al que cabría esperar, dada la condición leve en muchos casos y la baja tasa de hospitalizaciones (104, 105). Los elevados costes directos están relacionados con la alta frecuencia con que se presentan además de los consumos de medicamentos para su control (106). Más difícil resulta calcular los costes indirectos, es decir los derivados de la pérdida de días de trabajo o de la reducción de la productividad laboral (106). La evaluación de los costes de las distintas enfermedades tiene un gran interés para la sociedad y es necesaria para un mayor conocimiento de los problemas de salud, contribuyendo a establecer la equitativa distribución de recursos tanto humanos como materiales para atajar la repercusión de las enfermedades sobre los pacientes (107, 108, 109).

1.2. Epidemiología

La epidemiología es la disciplina científica que estudia la frecuencia y distribución de fenómenos relacionados con la salud y la enfermedad y sus determinantes en poblaciones específicas. También aplica este estudio al control de problemas de salud. Estudiar la distribución significa estudiar la medida de la frecuencia y variación de un fenómeno en grupos de población a lo largo del tiempo, en diferentes lugares o formados por diferentes tipos de personas (110).

El estudio incluye las investigaciones caracterizadas por la simple vigilancia y observación de fenómenos para medir su magnitud y sugerir hipótesis sobre su origen. Este tipo de investigaciones reciben el calificativo de descriptivas. Pero también incluye las investigaciones dirigidas a contrastar estas hipótesis mediante estudios observacionales y experimentales. Estas investigaciones reciben el nombre de analíticas. La epidemiología no sólo estudia enfermedades sino todo tipo de fenómenos relacionados con la salud, entre los que se encuentran causas de muerte como los accidentes o suicidios, hábitos de vida como el consumo de tabaco o la dieta y el uso de servicios de salud o la calidad de vida relacionada con la salud, entre otros. Los determinantes de estos fenómenos son todos los factores físicos, biológicos, sociales, culturales y de comportamiento que influyen sobre la salud (111).

Como todo método científico, el método epidemiológico tiene 3 fases. En primer lugar, observar o conocer la frecuencia y distribución de las enfermedades. A continuación, los patrones de frecuencia y distribución pueden dar lugar a hipótesis de asociación entre la enfermedad y un determinado factor. Finalmente, se diseña un estudio experimental para verificar o refutar la hipótesis planteada (112).

Los fenómenos relacionados con la salud y sus posibles determinantes dan lugar a algunas de las clasificaciones de las ramas de la epidemiología. Así, cuando el eje de clasificación son los fenómenos sanitarios surgen ramas como la epidemiología cardiovascular, del cáncer, o de los servicios sanitarios. Cuando el eje son los determinantes, surgen la epidemiología nutricional, laboral, o social. La epidemiología es una disciplina básica de la salud pública y de la medicina clínica, porque sus conocimientos pueden y deben ser aplicados al control de problemas de salud en ambos campos. Estas aplicaciones de la epidemiología permiten su

clasificación en dos ramas. La distinción entre la epidemiología general o de salud pública y la epidemiología clínica no estriba tanto en las técnicas utilizadas como en la porción de la historia natural de la enfermedad que es estudiada por cada una de ellas. La historia natural de una enfermedad es el conjunto de sucesos que van desde que un sujeto o grupo de sujetos resulta expuesto a las primeras causas de una enfermedad hasta que ésta se desarrolla y finalmente se resuelve con la curación total, la curación con secuelas o la muerte (113).

La epidemiología de salud pública estudia la primera parte de esta cadena de sucesos, es decir, la frecuencia y distribución de la enfermedad y sus determinantes, factores de riesgo o protección. Para ello se fija en sujetos sanos, generalmente viviendo en la comunidad, a los que sigue para observar como enferman (114).

La epidemiología clínica estudia la frecuencia y distribución de las consecuencias de la enfermedad y sus determinantes, los factores pronósticos. Para ello, suele fijarse en sujetos enfermos en los que mide posibles factores pronósticos y los sigue para observar la evolución de la enfermedad (115).

Sin embargo el rasgo que más diferencia a la epidemiología de otras disciplinas biológicas es el estudio de la frecuencia de fenómenos en grupos de sujetos, en poblaciones. Mientras que las ciencias experimentales en el laboratorio estudian, sobre todo, relaciones deterministas o no estocásticas, la epidemiología se centra en relaciones probabilísticas o estocásticas (116).

En epidemiología, los potenciales factores etiológicos pueden, en un caso concreto, producir o no producir una enfermedad. Así, una persona que fuma puede o no desarrollar un infarto agudo de miocardio, siendo la segunda posibilidad algo más frecuente que la primera. Nuestros factores causales aumentan la frecuencia o probabilidad de desarrollar una enfermedad -son factores de riesgo- pero no aseguran su desarrollo. Para poner de manifiesto su efecto se necesitan grupos de personas en los que medir la frecuencia de la enfermedad y compararla entre los expuestos y no expuestos al factor etiológico. En los estudios experimentales de laboratorio se suelen abordar relaciones deterministas. En ellas el factor causal es suficiente para producir su efecto. Siempre que alguien se exponga a él sufrirá su efecto. Para ponerlo de manifiesto bastará con estudiar un solo sujeto (117).

1.2.1. Clasificación y tipos de estudios epidemiológicos

Los estudios epidemiológicos se clasifican en descriptivos o analíticos. Los estudios descriptivos valoran la frecuencia y la distribución de las enfermedades en las poblaciones, con relación a las variables de persona, lugar y tiempo. Estos pueden aportar indicios que contribuyan a generar hipótesis sobre asociaciones entre factores de exposición y estados de salud o enfermedad (118). Los estudios analíticos valoran los determinantes de los estados de salud o enfermedad, comprobando o rechazando las hipótesis generadas por los estudios descriptivos, con el objetivo de identificar factores de riesgo o protectores de una enfermedad. Entre los estudios descriptivos encontramos los estudios ecológicos, realizados sobre poblaciones, y las series de casos y estudios transversales, realizados sobre individuos. Los estudios analíticos pueden ser observacionales (estudios de casos y controles, estudios de cohortes) o intervencionistas (ensayos clínicos, ensayos de campo, ensayos comunitarios) (119). La mejor evidencia científica al analizar la frecuencia de una enfermedad se obtiene a partir de censos o de estudios sobre muestras aleatorias de la población general, y los estudios sobre cohortes de casos incidentes ofrecen la mejor evidencia para describir la historia natural o el pronóstico. Independientemente del nivel de evidencia, cada tipo de estudio tiene sus ventajas y limitaciones, que los hacen adecuados para diferentes escenarios. Finalmente, hay dos aspectos claves que influyen en la calidad de un estudio epidemiológico: hay que definir claramente la población de referencia sobre la que se van a extrapolar los resultados del estudio (estudios de base poblacional o de base hospitalaria) y los casos deben ser correctamente definidos, con una adecuada selección de las variables por las que se va a definir un proceso y de las escalas de medida que se van a utilizar (120,121).

1.2.1.1. Estudios descriptivos

Estos estudios describen la frecuencia y las características más importantes de un problema de salud. Los datos proporcionados por estos estudios son esenciales para los administradores sanitarios así como para los epidemiólogos y los clínicos. Los primeros podrán identificar los grupos de población más vulnerables y distribuir los

recursos según dichas necesidades y para los segundos son el primer paso en la investigación de los determinantes de la enfermedad y la identificación de los factores de riesgo (122).

Los principales tipos de estudios descriptivos son: los estudios ecológicos, los estudios de series de casos y los transversales o de prevalencia (123).

-Estudios ecológicos: Estos estudios no utilizan la información del individuo de una forma aislada sino que utilizan datos agregados de toda la población (124). Describen la enfermedad en la población en relación a variables de interés como puede ser la edad, la utilización de servicios, el consumo de alimentos, de bebidas alcohólicas, de tabaco, la renta per cápita. Un ejemplo de este estudio sería correlacionar la mortalidad por enfermedad coronaria con el consumo per cápita de cigarrillos. Estos estudios son el primer paso en muchas ocasiones en la investigación de una posible relación entre una enfermedad y una exposición determinada. Su gran ventaja reside en que se realizan muy rápidamente, prácticamente sin coste y con información que suele estar disponible. Así por ejemplo, los datos demográficos y el consumo de diferentes productos se pueden correlacionar con la utilización de servicios sanitarios, con registros de mortalidad y registros de cáncer. La principal limitación de estos estudios es que no pueden determinar si existe una asociación entre una exposición y una enfermedad a nivel individual. La falacia ecológica consiste precisamente en obtener conclusiones inadecuadas a nivel individual basados en datos poblacionales. Otra gran limitación de los estudios ecológicos es la incapacidad para controlar por variables potencialmente confusoras. La asociación o correlación que encontremos entre dos variables puede ser debida a una tercera variable que a su vez esté asociada con la enfermedad y la exposición objeto de estudio.

-Series de casos: Estos estudios describen la experiencia de un paciente o un grupo de pacientes con un diagnóstico similar. En estos estudios frecuentemente se describe una característica de una enfermedad o de un paciente, que sirven para generar nuevas hipótesis. Muchas veces documentan la presencia de nuevas enfermedades o efectos adversos y en este sentido sirven para mantener una vigilancia epidemiológica.

Estos estudios aunque son muy útiles para formular hipótesis, no sirven para evaluar o testar la presencia de una asociación estadística. La presencia de una asociación puede ser un hecho fortuito. La gran limitación de este tipo de estudios es en definitiva la ausencia de un grupo control (125,126).

-Estudios transversales: El estudio transversal es un estudio observacional que mide tanto la exposición como el resultado en un punto determinado en el tiempo. Este diseño se lleva a cabo usualmente para estimar la prevalencia de una enfermedad dentro de una población específica, y es común que uno de sus propósitos centrales sea proporcionar información para efectuar intervenciones de salud pública. Hay dos tipos de estudios transversales, los de tipo descriptivo que caracterizan la prevalencia de una enfermedad en la población de interés, y los analíticos que examinan la relación entre la exposición y la enfermedad, aunque sin poder establecer juicios definitivos de causalidad debido a la ambigüedad temporal que ocurre al coleccionar simultáneamente información acerca de los factores de riesgo y el resultado de interés. Este tipo de estudios, estudian simultáneamente la exposición y la enfermedad en una población bien definida en un momento determinado. Esta medición simultánea no permite conocer la secuencia temporal de los acontecimientos y no es por tanto posible determinar si la exposición precedió a la enfermedad o viceversa (127).

La prevalencia es una medida de frecuencia especialmente importante para el médico clínico, debido a la influencia que ejerce en la sospecha de un diagnóstico determinado o para futuras investigaciones (128,129). Dado que regularmente el estudio transversal se ejecuta para estimar la prevalencia de un evento determinado en una población específica (fundamentalmente una enfermedad o un factor de riesgo), es común que uno de sus propósitos centrales sea proporcionar información para valorar el estado de salud de una comunidad y determinar sus necesidades, y posteriormente efectuar intervenciones de salud pública. Así mismo, sirven como todos los estudios descriptivos para formular hipótesis etiológicas (130,131).

La realización de este tipo de estudios requiere definir claramente: la población de referencia sobre la que se desea extrapolar los resultados, y la población susceptible de ser incluida en nuestra muestra delimitando claramente los que pueden ser incluidos en dicho estudio. También son importantes: la selección y definición de

variables por las que se va a caracterizar el proceso, y las escalas de medida a utilizar (132).

1.2.1.2. Estudios analíticos

-Estudio de casos y controles: Este tipo de estudio identifica a personas con una enfermedad (u otra variable de interés) que estudiemos y los compara con un grupo control apropiado que no tenga la enfermedad (133). La relación entre uno o varios factores relacionados con la enfermedad se examina comparando la frecuencia de exposición a éste u otros factores entre los casos y los controles. A este tipo de estudio que es de los más utilizados en la investigación se le podría describir como un procedimiento epidemiológico analítico, no experimental con un sentido retrospectivo, ya que partiendo del efecto, se estudian sus antecedentes, en el que se seleccionan dos grupos de sujetos llamados casos y controles según tengan o no la enfermedad. En los estudios de casos y controles tenemos casos expuestos, casos no expuestos, controles expuestos y controles no expuestos. En este estudio la frecuencia de exposición a la causa entre los casos se compara con la frecuencia de exposición en una muestra que represente a los individuos en los que el efecto no se ha producido (134,135).

-Estudio de cohortes (o de seguimiento): En este tipo de estudio los individuos son identificados en función de la presencia o ausencia de exposición a un determinado factor. En ese momento todos están libres de la enfermedad de interés y son seguidos durante un período de tiempo para observar la frecuencia de aparición del fenómeno que nos interesa. Si al finalizar el período de observación la incidencia de la enfermedad es mayor en el grupo de expuestos, podremos concluir que existe una asociación estadística entre la exposición a la variable y la incidencia de la enfermedad (136).

1.2.1.3. Estudios experimentales

En los estudios experimentales el investigador maneja las condiciones de la investigación. Este tipo de estudios se utilizan para evaluar la eficacia de diferentes

terapias, de actividades preventivas o para la evaluación de actividades de planificación y programación sanitarias. Como en los estudios de seguimiento, los individuos son identificados en base a su exposición, pero a diferencia de estos, en los estudios experimentales es el investigador el que decide la exposición. El gran control que se tiene sobre el diseño facilita la interpretación de las asociaciones como causales. Para el médico clínico es de gran interés poder realizar inferencias causales en medio de la incertidumbre que rodea la práctica clínica ya sea en actividades de prevención, de diagnóstico o terapéuticas. El ensayo clínico es el estudio experimental más frecuente. Los sujetos son pacientes y evalúa uno o más tratamientos para una enfermedad o proceso. La validez de este estudio radica fundamentalmente en que el proceso aleatorio haga los grupos comparables en las variables más relevantes en relación al problema a estudiar (136).

1.3. Estudios ecológicos

Los estudios ecológicos (o correlacionales) utilizan datos de poblaciones para comparar la frecuencia de enfermedad entre diversos grupos durante un mismo periodo, o en un mismo grupo durante periodos diferentes. Las variables de interés pueden ser la edad, el sexo y variables étnicas, socioeconómicas o geográficas. Su principal ventaja es que pueden realizarse rápidamente y con escasos costes, al basarse en información fácilmente disponible (censos, registros de mortalidad, registros de cáncer, datos de consumo de productos). Suelen ser el primer paso en la investigación de una posible relación entre una exposición y una enfermedad. La posibilidad de realizar un análisis de toda la población hace que sean útiles en el estudio de enfermedades raras. Las conclusiones de este tipo de diseño se refieren solo a las poblaciones y no a los individuos. No es posible extrapolarlas a la exposición del individuo y el riesgo de presentar la enfermedad (lo que se conoce como «falacia ecológica»). Los estudios ecológicos tampoco pueden controlar la presencia de variables de confusión.

Los estudios ecológicos con frecuencia son utilizados como un primer paso en la investigación de cierto tipo de exposición y su relación con la incidencia, prevalencia, mortalidad o cualquier otro factor de interés en una población (137).

En general, las investigaciones que se realizan con este tipo de diseño son rápidas y poco costosas debido a que utilizan datos que han sido recopilados de manera previa y rutinaria en censos, encuestas, programas de vigilancia epidemiológica, registros de enfermedades específicas y programas de control de la contaminación, entre otros. Gracias a esto, en vez de realizar un prolongado y costoso estudio donde se evalúe de manera individual la influencia de un determinado factor sobre la salud, se puede consultar la información nacional sanitaria y mostrar con poco dinero y tiempo si hay una correlación (138).

Los estudios ecológicos presentan lógicamente problemas metodológicos (139). Al igual que los estudios individuales, los ecológicos pueden presentar un efecto de confusión debido a que, por lo general, no se miden o evalúan todas las posibles variables que pueden alterar la asociación de interés. Desde un punto de vista de población, las características de una variable de confusión son las mismas que a nivel individual. Sin embargo, si el estudio ecológico explora una asociación individual, es preferible manejar la confusión mediante la incorporación de unidades de observación más pequeñas (pasar de analizar países a analizar municipios, por ejemplo).

El sesgo de migración es una manifestación frecuente del sesgo de selección en los estudios ecológicos, que se refiere a la movilización de individuos entre áreas geográficas (140). Este sesgo puede afectar la validez de los estudios ecológicos, especialmente en cuanto al tiempo de latencia de enfermedades crónicas. Las migraciones que tienen lugar durante el periodo de latencia dependerán de la edad, el sexo y, en particular, del tamaño físico de las unidades geográficas donde se presenta la exposición. Se debe considerar la posibilidad de que exista un sesgo de migración cuando las áreas residenciales o geográficas se utilizan como una medida sustituta de la exposición y se estudian los factores determinantes de la enfermedad, en particular con largos periodos de latencia.

Otra fuente de sesgo de selección es que se disponga de información solo para las unidades de observación con mejores condiciones, lo que restringiría la observación de las asociaciones exclusivamente a las mejores circunstancias, lo que impediría extrapolarlas e implicaría poca validez externa de los hallazgos.

Los estudios ecológicos pueden verse limitados por las fuentes de información, es decir, por la disponibilidad y la calidad de los datos, hasta el punto de comprometer la validez de sus resultados. Cuando no se disponga de datos o estos no sean de calidad, es preferible no llevar a cabo el estudio. Es importante señalar, no obstante, que en los últimos años se han desarrollado métodos para solventar problemas como la falta de información y los datos de poca calidad. Uno de los más utilizados es la generación de modelos de regresión para asignar las estimaciones a las unidades de observación que carecen de datos. También, se hacen ajustes mediante el subregistro de eventos con base en variables demográficas, como la edad y el sexo, y el ajuste frente a definiciones incongruentes, con lo que se busca eliminar eventos no relacionados con el indicador estimado. Los modelos de niveles múltiples también se han propuesto como alternativa para identificar problemas relacionados con la estimación de la variable o un indicador propuesto para diferentes niveles. En conclusión, los estudios ecológicos siguen siendo una fuente importante de información en salud ambiental, pese a que suelen ser poco valorados por los epidemiólogos. En el siglo XXI, la salud ambiental ya no fundamenta sus conocimientos exclusivamente en la toxicología y la epidemiología, lo que ha permitido el empleo de disciplinas emergentes, especialmente en las ciencias sociales y del comportamiento, cuyos modelos conceptuales y métodos de estudio amplían la gama de posibilidades metodológicas con enfoques cuantitativos y cualitativos (141).

El uso adecuado de los estudios ecológicos puede ayudar a superar el individualismo metodológico preponderante en la epidemiología y la salud ambiental, el cual postula que todos los fenómenos provienen de los individuos (entendidos como constituyentes últimos del mundo social) y se funda en la premisa de que así se tiene acceso directo a los hechos tanto en el ámbito individual como en el social, concibiendo el segundo como derivado del primero. Esta rígida postura ya ha sido cuestionada y muchos hallazgos han demostrado sus limitaciones. En este sentido, la investigación en salud ambiental está llamada a ser pionera, pues el estudio del contexto en el que los individuos y las poblaciones viven y se exponen a diferentes agentes, le es inherente: la comprensión de las condiciones de salud de los

individuos depende de la comprensión de dichas condiciones en las poblaciones (142).

En cuanto al tipo de variables que se manejan en este tipo de estudios podemos mencionar la siguientes:

-Las variables agregadas o compuestas son las más usuales en estudios epidemiológicos. Se caracterizan por tener una variable correspondiente en el nivel individual y suelen ser medidas de resumen de un grupo de individuos y expresarse como promedios, medianas o proporciones (143).

-Las variables globales corresponden al otro extremo del espectro con respecto a las variables agregadas: no se pueden medir a nivel individual, por lo que son variables ecológicas exclusivamente. Un ejemplo es el análisis de los anillos de los árboles para medir retrospectivamente cambios en el ambiente, como incendios, inundaciones y otras condiciones climáticas. Estas medidas no tienen un referente individual humano, pues son atributos específicos del ambiente físico (144).

-Las variables ambientales se encuentran en un nivel intermedio entre las variables agregadas y las globales. Suelen ser atributos contextuales que tienen efecto sobre la salud de las poblaciones y pueden medirse a nivel ecológico o individual, aunque por razones logísticas generalmente se miden a nivel ecológico. Vistas desde la óptica individual, estas variables tienen menor heterogeneidad, lo que puede llevar a negar asociaciones reales. Un ejemplo de este tipo de variables se da en los estudios que relacionan contaminación atmosférica y salud, en los cuales se usan medidas recolectadas mediante redes de vigilancia cuyos hallazgos permiten el establecimiento de amplias categorías de exposición (145).

Sea cual sea el tipo de variable empleada, los estudios siempre emplean dos puntos de vista matemáticos en cuanto a las variables.

Los estudios cuantitativos suelen enfocarse en factores de las unidades de observación que se asocian con un resultado numérico. Cuando la variable es cuantitativa se presenta el problema fundamental de que en dos o más grupos la medida de tendencia central puede ser igual, aunque la dispersión sea diferente; en estos casos puede usarse el coeficiente de variación, o el coeficiente de asimetría; el

segundo es preferible cuando no se indica una medida de la precisión relativa, pero siempre debe tenerse precaución con las inferencias resultantes. Se pueden usar muchas otras formas de medición más complejas desde el punto de vista matemático, aunque siempre deben adecuarse a los objetivos específicos de cada estudio (146).

Esto contrasta con los estudios cualitativos, en los que las categorías son el centro del análisis. Una categoría puede definirse como un conjunto de contenidos que tienen algo en común. Se considera que las categorías deberían ser exhaustivas y mutuamente excluyentes (147).

1.3.1. Clasificación de los estudios ecológicos.

Los estudios epidemiológicos ecológicos pueden clasificarse según el tipo de inferencias, en exploratorios y analíticos, y en cada una de estas categorías pueden llevarse a cabo estudios en varios grupos y series temporales (148).

-Los estudios exploratorios consideran las diferencias geográficas en la aparición de un evento en varios conjuntos, generalmente regiones. Su objetivo es la búsqueda de patrones en la distribución de la morbilidad (incidencia o prevalencia) o de la mortalidad, que sugieran hipótesis etiológicas; en estos estudios no se suelen medir las exposiciones y, en general, se hacen análisis estadísticos. La práctica general indica que deben usarse datos estandarizados por variables como la edad o el sexo, que pueden ser variables de confusión poderosas para las asociaciones causales que pretenden explorarse (149).

-Los estudios analíticos analizan hipótesis de asociación entre exposiciones ambientales y eventos en salud, por lo cual se suelen utilizar análisis estadísticos más avanzados, como las regresiones múltiples, para ajustar las asociaciones según variables de confusión (150).

-Los estudios de series de tiempo son más complejos e incorporan métodos estadísticos que buscan predecir eventos con base en los datos previos, por lo que uno de los objetivos más importantes es identificar el desfase temporal que mejor predice un evento futuro basado en los anteriores (151). Los estudios de caso único suelen ser análisis de casos o estudios de series de tiempo. Los primeros son, por lo general, estrictamente descriptivos.

Los estudios ecológicos de epidemiología espacial incorporan la georreferenciación de los lugares de obtención de los datos. Los análisis de georreferenciación se perfeccionan en la misma medida que lo hacen las nuevas tecnologías, capaces de evaluar y ponderar efectos ajustados según diferentes variables ambientales, como en el caso de los niveles de exposición en diferentes áreas geográficas mediante, por ejemplo, los modelos geográficamente ponderados (152). En ellos se describen y analizan las variaciones geográficas de un evento en salud con respecto a agentes ambientales o de otra índole. Existen cuatro tipos de estudios en epidemiología espacial: El mapeo de enfermedades (utilizado en estudios descriptivos que se basan en características geográficas); las correlaciones geográficas (catalogadas como estudios ecológicos que se basan en la comparación de unidades geográficas) (153); la detección de grupos, es decir, agregaciones de enfermedades similares o relacionadas en grupos específicos de una población (154), y, por último, los estudios de fuentes puntuales que pueden ampliar la evidencia sobre la causalidad (cuando, por ejemplo, un patrón de exposición y respuesta puede demostrarse en relación con un punto, línea o área de fuentes) (155).

1.4. Epidemiología de la alergia

Los estudios de prevalencia de alergia son muy escasos en la población general española. Por tanto, se deberían realizar estudios en nuestro país para establecer la prevalencia de enfermedades alérgicas en una muestra de población general, así como establecer las causas y las manifestaciones de esta alergia. También es importante verificar, en nuestra zona, estas mismas necesidades y si tenemos datos acerca de la idea de que la alergia se comporta como una epidemia y afecta a una proporción importante de la población. En los estudios efectuados hasta ahora, más importantes y con mayor proporción de pacientes, por ejemplo en los sucesivos Alergológica (156, 157, 158) realizados por la Sociedad Española de Alergología e Inmunología Clínica, en los cuales se muestra el panorama actualizado de la asistencia alergológica, no se ha podido estimar la prevalencia poblacional de alergia por tratarse de muestreos no representativos (muestra que acudía al alergólogo con la sospecha de padecer una enfermedad de origen alérgico). No existen estudios que

tengan en cuenta los datos agregados que se obtienen de estudios ecológicos y que pone el punto de mira en las poblaciones y no en los individuos.

La alergología es una especialidad clínica de demanda asistencial cambiante en cuanto a prevalencia de ciertos trastornos, sensibilización social creciente frente a los problemas alérgicos, cambios sociosanitarios en la población y cambios en la oferta del sistema sanitario español, novedades científico-técnicas de reciente introducción en la especialidad, novedades del arsenal terapéutico antialérgico, entre otros, siendo todos ellos factores que han cambiado las características de la práctica alergológica en los últimos años. El objetivo general del último estudio Alergológica 2015 fue obtener información de la práctica clínica alergológica española sobre los casos de pacientes que consultan por primera vez con un alergólogo durante el año 2014 y primer trimestre del 2015, e identificar las modificaciones relevantes acaecidas en la misma durante la última década, replicando los objetivos centrales de ediciones previas de Alergológica realizadas en los años 1992 y 2005. La información recogida en este estudio ha sido de utilidad para valorar el estado de salud de la comunidad en relación con las enfermedades alérgicas y para determinar sus necesidades. Así mismo, se intentaba que sirviese como todos los estudios descriptivos para formular hipótesis etiológicas. Con este estudio se buscaba analizar las características epidemiológicas y clínicas de la población que acudía por primera vez a las consultas de alergología españolas, y describir la prevalencia de los distintos problemas alérgicos atendidos (rinitis y conjuntivitis, asma bronquial, urticaria y angioedema, dermatitis atópica, dermatitis de contacto, alergia alimentaria, alergia a medicamentos, hipersensibilidad a insectos, falsa alergia, etc). También analizar las posibles diferencias actuales de distribución regional de los trastornos alérgicos en España, junto con la posible variabilidad estacional de la demanda de cuidados alergológicos en España. Así mismo, describir el proceder diagnóstico del especialista y el uso actual de tecnología diagnóstica en las consultas alergológicas, las pautas de tratamiento actuales y las medidas de prevención instauradas en las consultas alergológicas. Finalmente, estimar la repercusión sociosanitaria de las enfermedades alérgicas en la calidad de vida del paciente y en su actividad escolar o laboral.

El contraste de los resultados conocidos de los estudios Alergológica 1992 y Alergológica 2005 con la información comparable y actualizada que generó este estudio, intentaba permitir identificar modificaciones relevantes y sugerir líneas potenciales de desarrollo profesional, mejora asistencial e investigación clínica.

Se trataba de un estudio epidemiológico observacional descriptivo de tipo transversal, con recogida prospectiva de información sobre pacientes atendidos por primera vez en una amplia muestra de consultas alergológicas estratificadas geográficamente en todo el territorio nacional. En cuanto al análisis estadístico, se estimaron las prevalencias de los distintos trastornos alérgicos, y se describieron las restantes variables cualitativas del estudio mediante el cálculo de frecuencias relativas (%) y del intervalo de confianza al 95%. La descripción de variables cuantitativas se efectuó mediante las medidas de centralización y dispersión al uso (media y desviación típica), empleando algunos indicadores de posición característicos (mediana, cuartiles) rango intercuartílico y valores adyacentes superior e inferior en los casos de dispersión amplia o atípica de los datos. Se compararon diferencias entre porcentajes de variables cualitativas, determinando el intervalo de confianza al 95% de la diferencia de porcentajes apreciados y empleando la prueba estadística de Fisher, según requería cada circunstancia. La comparación de valores promedio en variables cuantitativas se efectuó mediante la prueba T (o U de Mann-Whitney, si las condiciones de las submuestras así lo exigían) o el análisis de la varianza (o prueba de Kruskal-Wallis) en los casos de comparaciones múltiples, complementado por un test *a posteriori* de comparación múltiple (Bonferroni, Tukey). Para detectar la posible existencia de características (personales, clínicas, asistenciales o profesionales), predictoras de circunstancias de interés clave para los investigadores (estancia prolongada en lista de espera, uso de cierta tecnología diagnóstica o terapéutica, etc), se diseñaron modelos de regresión logística utilizando como variable de respuesta binaria la cuestión objeto de interés. Se detectaron los *ítems* que expresan un valor predictivo independiente, determinando su correspondiente intervalo de confianza al 95%.

En resumen, los resultados mostraron que 6 de cada 10 motivos (62%) de primera consulta al alergólogo son por rinoconjuntivitis. El segundo motivo de consulta es el asma (23,4%) y el tercero la hipersensibilidad a los medicamentos (17,7%). En

comparación con el informe anterior Alergológica 2005 han aumentado las consultas por rinitis (62% *versus* 55%) y las de alergia a alimentos (10% *versus* 7%) y se han reducido las consultas por asma (23% *versus* 28%). La media del tiempo de espera para una primera consulta, a nivel global, fue de 45 días. La rinoconjuntivitis ha sido el principal motivo de consulta en Alergológica 2015 y muchos pacientes que se atienden por primera vez en las consultas de alergología acuden por presentar rinoconjuntivitis, dato que refuerza la importancia de este trastorno (en lo que a frecuencia se refiere) en el seno de esta especialidad médica. Las cifras de rinoconjuntivitis duplican a las de asma bronquial (21%), segundo motivo de consulta en esta edición de Alergológica (28% en Alergológica 2005). Es importante reseñar que un 17,5% de los pacientes consulta a la vez por ambas enfermedades. En total, 618 pacientes de la muestra estudiada presentaban asma bronquial, lo que representa un 21,2% de los sujetos incluidos. Al igual que en las ediciones previas, el asma se sitúa en la segunda enfermedad más frecuente estudiada en las consultas de alergología, tras la rinitis/conjuntivitis, que es con diferencia la enfermedad más frecuente (62%; 54,7% en Alergológica 2005). No obstante, la prevalencia de asma es notablemente inferior a la observada en los estudios previos, que era del 35% en Alergológica 1992 y 28% en Alergológica 2005.

Del total de pacientes que fueron atendidos en consultas de alergia, trescientos treinta y seis pacientes (11,5%) fueron diagnosticados de urticaria con o sin angioedema (urticaria/ angioedema). Datos similares se obtuvieron en el Alergológica 2005 con una prevalencia del 11%. Sigue representando el cuarto motivo de consulta más frecuente por detrás de la rinitis, el asma y la alergia a los medicamentos.

El número de pacientes que acudió al alergólogo por dermatitis atópica fue de 98 (un 60% de niños y un 40% de adultos), es decir, solo un 3,4% (2,8–4,1%) de toda la muestra, porcentaje igual al observado en Alergológica 2005 (3,4%, 2,89%–3,90%) con un predominio de niños (63,9% de niños y 36,1% de adultos). Si estratificamos la muestra global por rangos de edad, la mayor parte de los afectados se sitúan entre los 0 y 5 años (19,1%), el 10% en el grupo de 6-15 años y el 1,7% en el de mayores de 15 años.

Se diagnosticaron 124 casos de dermatitis de contacto, lo que supone un 4,3% del total de la muestra. La prevalencia de la dermatitis de contacto ha resultado similar a la del anterior estudio de Alergológica 2005, que fue del 4,2%, y superior a la de Alergológica 1992, del 2,4%. Al igual que sucedió en las ediciones previas, la prevalencia de la dermatitis de contacto ha sido mayor que la de la dermatitis atópica (3,4%), y menor que la de la urticarial/angioedema (11,5%). Esta cifra del 4,2% de pacientes diagnosticados podría considerarse inferior a lo esperado, si tenemos en cuenta que la prevalencia de la alergia de contacto (sensibilización frente a algún alérgeno de contacto) en la población general se encuentra en torno al 20%, o incluso el 27% en los estudios multicéntricos europeos más recientes y que hasta un 8,2% de la población general refiere padecer esta patología diagnosticada por un médico, y que entre las dermatitis, la alérgica es el motivo de consulta más frecuente en las consultas dermatológicas.

Se estima que la alergia a los alimentos afecta del 1% al 3% de la población general y que es más frecuente en los niños menores de tres años, en los que la prevalencia puede llegar hasta el 8%. Sin embargo, las diferencias en el diseño de los estudios y en los criterios diagnósticos dificultan las comparaciones.

En Alergológica 2015 la prevalencia de alergia a los alimentos en España entre los pacientes que acuden por primera vez a un alergólogo es del 11,4%, 331 casos) (en Alergológica 2005, 7,4%; 369 casos). Se mantiene en el mismo puesto de orden respecto a 2005, siendo la quinta enfermedad en orden de frecuencia de las diagnosticadas por los alergólogos en nuestro país, después de la rinoconjuntivitis, el asma, la hipersensibilidad a los medicamentos y la urticaria/angioedema. Hay una percepción general de que la frecuencia y gravedad de la alergia a los alimentos está aumentando. Sin embargo, los estudios poblacionales para estimar la prevalencia de una determinada patología son difíciles de realizar y hasta el momento no se ha podido demostrar de una forma inequívoca este aumento de la frecuencia de esta patología. El estudio Alergológica realizado con la misma metodología en sus tres ediciones Alergológica 1992, 2005 y 2015, fue muy útil para obtener tendencias de frecuencia de enfermedad. Comparándolos hemos constatado que la alergia a los alimentos se ha duplicado en nuestro país en poco más de un decenio, pasando de una prevalencia de 3,6% en 1992 al 7,4% en 2005 y al 11,4% en 2015.

En la última edición de Alergológica, 544 pacientes (735 en Alergológica 2005) consultaron por una posible alergia a medicamentos, lo que representa el 18,7% (14,7% en Alergológica 2005) de los pacientes incluidos en el estudio. Esta proporción mantiene el aumento progresivo detectado desde Alergológica 1992 donde esta cifra representaba el 12,63%. La alergia a los medicamentos constituye, al igual que en las anteriores ediciones del estudio Alergológica, la tercera causa de consulta en las unidades de alergología, tras la rinoconjuntivitis y el asma bronquial. En cuanto a la alergia a picaduras de los insectos, en esta muestra se diagnosticaron 73 pacientes, 2,5% (77 pacientes, 1,5% en Alergológica 2005) hipersensibles a las picaduras de los insectos. Se ve un aumento paulatino con respecto a Alergológica 1992 (0,7%). Hay un predominio de los pacientes del sexo masculino y de 20 a 50 años de edad.

El grupo de otras enfermedades no alérgicas supone el 3,9% (5,8% en Alergológica 2005) del total, 113 (290 en Alergológica 2005) pacientes de los 2.914 (4.991 en Alergológica 2005) del estudio, mientras que el diagnóstico de otras enfermedades alérgicas afecta a 104 pacientes, un 3,6% del total (45 pacientes, un 0,29% del total en Alergológica 2005). Se diagnosticaron otras enfermedades alérgicas (vasculitis, aspergilosis broncopulmonar alérgica, neumonitis alérgica extrínseca, eosinofilia pulmonar, infiltrados pulmonares por fármacos, anafilaxia idiopática, inmunodeficiencia, tos no filiada, urticaria pigmentosa, eritema nodoso, anisakiasis gastroalérgica, bronquitis eosinofílica) en 104 pacientes lo que supone un 3,6% de todos los pacientes del estudio (45 pacientes, un 0,9% en Alergológica 2005). En el estudio Alergológica 1992 se contabilizaron 76 pacientes.

En otro de los estudios de epidemiología alérgica realizado en España (159) en 2004, se admite que en los últimos decenios se está registrando un aumento de las enfermedades alérgicas en todo el mundo. Sin embargo, hay pocos estudios de prevalencia de alergia en la población general española. En este trabajo se intentaba establecer la prevalencia de enfermedades alérgicas en una muestra de población general adulta del estado español, así como establecer las causas y las manifestaciones de esta alergia. Se realizó un estudio observacional de tipo transversal con un tamaño de muestra que fue de 5.003 sujetos. La muestra se

seleccionó de forma aleatoria y automática del listín de teléfono, aplicando el control de cuotas de zona, el número de habitantes, el sexo y la edad. La recogida de datos se efectuó mediante encuesta telefónica siguiendo la técnica CATI (*Computer-assisted Telephone Interview technique*). Los resultados mostraron: de los 4.949 sujetos incluidos en el análisis, un 21,6% refirieron ser alérgicos.

La prevalencia fue mayor en mujeres (24,6%), en el grupo de 18 a 24 años de edad (26,9%) y en poblaciones de más de 500.000 habitantes (24,3%), mientras que fue menor en los varones (18,3%), en la región norte-centro (17,8%) y en poblaciones con menos de 10.000 habitantes (18,7%).

La rinoconjuntivitis fue la manifestación alérgica más frecuente (45,4%) seguida del asma bronquial (24,9%), la urticaria (24,6%), la dermatitis (21,5%) y el angioedema (6%). Las tres causas más frecuentes de alergia fueron con diferencia los pólenes (31,5%), los medicamentos (29,4%) y los ácaros del polvo doméstico (25,3%). Mucho menos frecuentes fueron los animales (6,8%), los metales (4,9%), los alimentos (4,8%), los hongos (3%), las picaduras de insectos (2,5%), la exposición al sol (1,9%) y el látex (0,8%).

No hubo diferencias significativas en relación al nivel socioeconómico.

Las conclusiones que obtuvieron son que la prevalencia de alergia en la población adulta española es del 21,6%, y es más frecuente en las mujeres y en los núcleos urbanos de más de 500.000 habitantes. La manifestación más frecuente es la rinoconjuntivitis. Las manifestaciones cutáneas de la alergia no son despreciables. Las causas más frecuentes de alergia son los aeroalérgenos (pólenes y ácaros) y los medicamentos.

Si tenemos en cuenta solo la población infantil, hay que destacar al estudio ISAAC (*International Study of Asthma and Allergies in Childhood*), el estudio internacional sobre asma y alergias en la infancia, que ha sido uno de los trabajos epidemiológicos más intensivos a nivel mundial (160). El Programa ISAAC finalizó formalmente en diciembre de 2012. ISAAC, el Estudio Internacional de Asma y Alergia en la Infancia, se formó en 1991 para facilitar la investigación sobre el asma, la rinitis alérgica y el eccema mediante la promoción de una metodología estandarizada capaz de ser utilizada en diversos lugares de todo el mundo. Ha sido un programa de

investigación epidemiológica mundial único establecido en 1991 para investigar el asma, la rinitis y el eccema en niños debido a la considerable preocupación de que estas condiciones estaban aumentando en los países occidentales y en desarrollo. ISAAC se ha convertido en el mayor proyecto de investigación colaborativa a nivel mundial jamás emprendido, en el que participaron más de 100 países y casi 2 millones de niños, y su objetivo es desarrollar medidas ambientales y de vigilancia de enfermedades con el fin de formar la base de futuras intervenciones para reducir la carga de enfermedades alérgicas y no alérgicas, especialmente en los niños de los países en desarrollo.

ISAAC se desarrolló a partir de una fusión de dos proyectos de colaboración multinacionales, cada uno de los cuales investigaba las variaciones en el asma infantil a nivel poblacional. Estas fueron una iniciativa de Auckland, Nueva Zelanda para llevar a cabo un estudio comparativo internacional de la gravedad del asma, y una iniciativa de Bochum, Alemania, para llevar a cabo un estudio internacional para controlar las tendencias temporales y los determinantes de la prevalencia del asma y la alergia en los niños (160).

ISAAC fue concebido para desarrollarse en tres fases distintas. La fase I se llevó a cabo en la mayoría de los países en el período comprendido entre 1994 y 1996. Consistió fundamentalmente en la exploración mediante un cuestionario estandarizado muy elemental de los síntomas relacionados con el asma en dos grupos distintos de edad: 13-14 y 6-7 años. Los cuestionarios dirigidos a los niños mayores fueron autocumplimentados; y los dirigidos a los niños pequeños fueron rellenados por los padres. En España, la fase I fue seguida por 10 centros distintos, y participaron en ella 30.626 niños de 13-14 años y 19.010 de 6-7 años.

Los objetivos de la fase I de ISAAC fueron describir la prevalencia y la gravedad del asma, la rinitis y el eccema en niños que viven en diferentes centros y hacer comparaciones dentro de los países y entre ellos. También obtener medidas de referencia para la evaluación de las tendencias futuras en la prevalencia y gravedad de estas enfermedades; y proporcionar un marco para una mayor investigación etiológica sobre el estilo de vida, el medioambiente, la genética y los factores de atención médica que afectan a estas enfermedades.

Los objetivos de la segunda fase de ISAAC fueron describir la prevalencia de marcadores de asma y alergia en niños que viven en diferentes centros, y hacer comparaciones dentro y entre los centros. También evaluar la relación entre la prevalencia de marcadores de asma y alergia y la prevalencia de síntomas de estas afecciones en niños que viven en diferentes centros. Además, se buscó estimar en qué medida la variación en la prevalencia y gravedad del asma y la alergia en los niños entre centros puede explicarse por diferencias en los factores de riesgo conocidos o sospechosos o por diferencias en el manejo de la enfermedad. Finalmente se exploraron nuevas hipótesis etiológicas sobre el desarrollo de asma y alergia en niños.

Los objetivos de la tercera fase de ISAAC fueron examinar las tendencias temporales en la prevalencia de asma, rinoconjuntivitis alérgica y eccema atópico en centros y países que participaron en la fase uno de ISAAC. También describir la prevalencia y gravedad del asma, la rinoconjuntivitis alérgica y el eccema atópico en centros y países que no participaron en la fase I y, además, examinar las hipótesis a nivel individual que han sido sugeridas por los hallazgos de la fase I, los análisis ecológicos posteriores y los avances recientes.

Dado que la creación de ISAAC estuvo motivada por la preocupación de un incremento de asma en países desarrollados este fue el primer dato que se investigó. Se encontró que los síntomas de asma, rinitis y eccema eran más comunes en algunos países occidentales de altos ingresos como el Reino Unido, Nueva Zelanda, Australia, pero no tan altos en otros como España (160, 161, 162). Además, algunos países de ingresos bajos y medios tenían valores de prevalencia para los síntomas de asma, rinitis y eccema que estaban al mismo nivel que algunos países occidentales de ingresos altos. Hubo variaciones notables en la prevalencia de los síntomas del asma y las enfermedades alérgicas en todo el mundo. Aunque la prevalencia de los síntomas de asma, rinitis y eccema son en general más bajos en los países en desarrollo, algunos países en desarrollo tienen niveles particularmente altos de estas condiciones y síntomas proporcionalmente más graves. La parte medular de este estudio se llevó a cabo en tres fases. La primera fase, efectuada en más de 721.000 niños y adolescentes demostró que existía una gran variabilidad en

la prevalencia del asma. Considerando solo respuestas positivas a la pregunta de ¿alguna vez ha tenido asma?, los países con menor prevalencia de esta enfermedad en niños fueron: Albania, Austria, Bélgica, Estonia, Alemania, India, Irán, Latvia, Polonia y Georgia (1,4 a 4,2%) y los de mayor prevalencia fueron: Australia, Costa Rica y Nueva Zelanda (26,5 a 27,1%). En adolescentes, los países con baja prevalencia fueron: Albania, Estonia, Etiopía, Indonesia, Irán, Polonia, Rusia, Corea del Sur y Uzbekistán (1,6 a 3%), y con alta prevalencia: Australia, Nueva Zelanda, Omán, Perú, Singapur y Reino Unido (20,7 a 28,2%).

Los objetivos de la segunda fase de ISAAC fueron describir la prevalencia de marcadores de asma y alergia en niños que viven en diferentes centros, y hacer comparaciones dentro y entre los centros y también evaluar la relación entre la prevalencia de marcadores de asma y alergia. Se realizó análisis de las interrelaciones de las pruebas de punción cutánea, IgE total y específica de alérgenos en centros de fase dos con diversas condiciones de vida (163). Básicamente consistía en obtener un cuestionario muy pormenorizado en una muestra de 1.000 niños normales por centro, a los que también se le realiza un prick con una batería de alérgenos estándar. A una submuestra de estos niños se les extrajo sangre -para IgE total, específica y estudios genéticos-, se les realizó una prueba de provocación bronquial y se analizó el polvo de su casa. Participaron 30 centros de 22 países. No hubo correlación clara entre sensibilización atópica y asma en ese momento y variaba según el país. Sin embargo, el grupo que mostraba relación entre atopía y asma se correlacionaba fuertemente con el tamaño del producto interior bruto (PIB) del país.

También los síntomas de la rinoconjuntivitis (rinitis con prurito y lagrimeo en los ojos) fueron comunes en centros de varias regiones (164). Los síntomas graves de rinoconjuntivitis se encontraron principalmente en los centros de países de ingresos medios y bajos, particularmente en África y América Latina ilustrando que esta condición es importante y puede causar una morbilidad significativa. Existen amplias variaciones globales en la prevalencia de los síntomas actuales de rinoconjuntivitis, siendo más altos en los países de ingresos altos frente a los de

bajos ingresos, pero la prevalencia de síntomas graves fue mayor en los países menos ricos. La comorbilidad con el asma es alta, particularmente en África, América del Norte y Oceanía. Este mapa global de prevalencia de síntomas es de importancia clínica para los profesionales de la salud. La prevalencia general promedio de los síntomas de rinoconjuntivitis fue del 14,6% para los niños de 13 a 14 años (rango 1-45%). Se observó variación en la prevalencia de síntomas de rinoconjuntivitis severa entre centros (rango 0-5,1%) y regiones (rango 0,4% en Europa occidental a 2,3% en África), observándose la mayor prevalencia principalmente en los centros de ingresos medios y bajos. La comorbilidad con asma y eccema varió del 1,6% en el subcontinente indio al 4,7% en América del Norte. Para los niños de 6 a 7 años, la prevalencia promedio de síntomas de rinoconjuntivitis fue del 8,5% y también se observaron grandes variaciones en la prevalencia de síntomas entre regiones, países y centros.

Las grandes variaciones mundiales en la prevalencia del asma encontradas en la fase I de ISAAC, incluso entre personas de origen genético similar que viven en diferentes entornos, nos llevaron a creer que los factores ambientales en lugar de los factores genéticos eran la causa de estas grandes variaciones. Las influencias genéticas se exploraron directamente en la fase dos (165, 166). Se detectaron asociaciones significativas con sibilancias en solo cuatro genes, y las variantes de solo dos de estos también se relacionaron con la inmunoglobina E (IgE) específica del alérgeno. También hubo asociaciones altamente significativas entre las variantes de SPINK5 y el eccema visible y entre las variantes de interleucina 13 y la IgE total. Estos hallazgos sugieren que, a pesar de la plausibilidad biológica de los mecanismos relacionados con IgE en el asma, la evidencia genética de esta vía todavía es escasa.

Durante mucho tiempo se ha creído que la alergia era la causa del asma, la rinitis y los síntomas del eccema. Sin embargo, en ISAAC fase dos se encontró una relación muy débil entre alergia (atopia) y asma (167). Se ha detectado una relación más débil de lo que se pensaba entre alergia y asma/rinitis. Aunque si que es verdad que la asociación de la atopia con cada una de estas enfermedades fue más fuerte en los centros más ricos que en los centros menos ricos (168, 169).

Se realizaron análisis ecológicos para detectar los factores comunes que contribuyeran al aumento de la prevalencia del asma. Los primeros informes de prevalencia de asma encontraron altas prevalencias en países ricos, y hubo informes anecdóticos de bajas prevalencias en países no ricos. Esto llevó a explorar la relación entre el producto nacional bruto (PNB) y los síntomas, pero hubo una débil relación (170). No hubo correlación clara entre sensibilización atópica y asma en ese momento y variaba según el país. Sin embargo el grupo que mostraba relación entre atopía y asma se correlacionaba fuertemente con el tamaño del PIB del país.

Existe una amplia evidencia de estudios a nivel individual de que la contaminación del aire puede agravar el asma existente. Sin embargo, la pregunta es si la contaminación del aire influye en la proporción de niños que tienen síntomas de asma en una población. Se encontró una relación inversa débil entre la contaminación del aire por partículas a nivel de ciudad (PM10) y los síntomas de las tres afecciones, incluso después de controlar el PNB, que tiene una fuerte asociación inversa con la contaminación del aire (171).

Un corolario de la hipótesis de la higiene es que el uso de antibióticos puede aumentar el riesgo de asma al reducir el efecto protector de la exposición microbiana, por ejemplo, a través de la interrupción de la microbiota intestinal normal (172). Pero en los datos en 28 países utilizando ventas de antibióticos, las relaciones entre la prevalencia de los síntomas y la exposición a los antibióticos no fueron claras: se encontró una mezcla de efectos inversos y positivos débiles entre las prevalencias de los síntomas y las ventas totales de antibióticos y las ventas de antibióticos de amplio espectro. Este análisis sugirió que incluso si había una posible asociación causal del uso de antibióticos con el riesgo de asma, no parecía explicar las diferencias mundiales entre los países.

Como el clima afecta a poblaciones enteras, los estudios ecológicos son ideales para examinar la relación entre la prevalencia de enfermedades y las condiciones climáticas entre las poblaciones. En los análisis mundiales se observaron pocas asociaciones significativas (173). A medida que el mundo se ve más afectado por el

cambio climático, puede haber algunas regiones como Europa occidental donde la prevalencia de enfermedades se ve afectada por factores potencialmente modificables, como la humedad y la temperatura, pero a nivel mundial nuestros análisis ecológicos mostraron poco efecto.

Los patrones dietéticos han cambiado rápidamente con la modernización u occidentalización, y el alejamiento asociado de los alimentos de origen vegetal y la adición de grasas artificiales podrían afectar la prevalencia de los síntomas. No se encontraron asociaciones con el consumo de la carne y la leche, pero hubo un patrón de asociación inversa entre los alimentos de origen vegetal y los síntomas de las tres afecciones (174). El análisis en los países europeos de los ácidos grasos trans encontró una asociación positiva, lo que sugiere que las grasas artificiales pueden ser un factor en la prevalencia de las tres condiciones. Por lo tanto, las influencias dietéticas en las tres condiciones requieren más investigación.

Los estudios han encontrado que la prevalencia de los síntomas de la polinosis y el asma tiende a ser menor en las zonas rurales que en las urbanas, y más baja entre las personas que viven en granjas, donde es probable que haya una mayor exposición al polen. En el análisis ecológico de fase uno, la exposición al polen alergénico se evaluó mediante exposiciones alrededor de las fechas de vida temprana (175). No pareció aumentar el riesgo de adquirir síntomas de alergia respiratoria, e incluso puede dar cierta protección, pero esto no se ha encontrado consistentemente fuera de Europa y Estados Unidos. El grado de consistencia en las asociaciones inversas sugiere la posibilidad de un efecto protector del polen sobre la alergia.

Dada la fuerte relación entre el tabaco y los síntomas de asma a nivel individual, se investigó este dato a nivel colectivo y no se detectó asociación entre asma y consumo de tabaco (176). Sin embargo, generalmente hubo una relación positiva entre las mujeres que fumaban, pero una asociación inversa entre los hombres que fumaban y las tres afecciones. Este análisis indicó que la asociación a nivel individual bien

establecida entre el tabaquismo de los padres y el asma no tuvo en cuenta las diferencias internacionales en la prevalencia del asma.

En cuanto a la relación de la tuberculosis con asma (177, 178). Hubo informes mixtos sobre si la inmunización no tuvo ningún efecto sobre el asma, la rinitis y el eccema atópico, o fue potencialmente un factor de riesgo relacionado con las hipótesis de higiene. Los análisis a nivel de país no mostraron asociaciones. Los análisis a nivel central más potentes mostraron pequeñas relaciones inversas entre las vacunas DTP (difteria, tétanos, tos ferina) solo en el grupo de mayor edad, sin asociaciones con vacunas BCG (bacilo de Calmette-Guérin causante de la tuberculosis). En vista de informes anteriores de que la inmunización podría ser un factor de riesgo para el asma, este resultado principalmente nulo es tranquilizador para los programas de inmunización de la población, dada su importancia para la salud infantil.

Los últimos resultados del estudio ISAAC (fase 3) muestran que el asma presenta estabilización o descenso y que ha ocurrido principalmente en los países desarrollados, mientras que en los países en desarrollo continúa aumentando. Las causas de estos cambios aún son desconocidas. Una de las explicaciones con más aceptación es que factores ambientales, en especial motivados por la «occidentalización» de la forma de vida, están modificando la respuesta inmunológica y se especula que la etapa prenatal o los primeros años de vida son los periodos más susceptibles para que ocurran estos cambios (179). En este contexto, una de las teorías más aceptadas, denominada «hipótesis de la higiene», postula que la ausencia de la exposición habitual a microorganismos hace que el sistema inmunológico tenga una inmadurez relativa y que, por lo tanto, sus respuestas no sean de tipo protector (promovidas por la población Th1 de células T) sino de tipo alérgico (promovidas por la población Th2 de células T). Otras hipótesis están relacionadas con exposiciones a diversos factores como humo de tabaco, alérgenos, contaminantes atmosféricos, contaminantes intramuros o dieta de tipo occidental. Un mecanismo común a través del cual algunos de estos factores podrían estar operando son las modificaciones epigenéticas como la metilación del ADN o la

acetilación de histonas, que promueven silenciamiento o cambios en la expresión de genes, sin implicar cambios en la secuencia del ADN.

A raíz de los informes de los países de habla inglesa en el decenio de 1990 sobre los aumentos en la prevalencia del asma a partir del decenio de 1980, se esperaban aumentos continuos de la prevalencia. Sin embargo, ISAAC encontró que en la mayoría de los países de alta prevalencia, particularmente en los países de lengua inglesa, la prevalencia de los síntomas del asma cambió poco entre la fase uno y la fase tres, e incluso disminuyó en algunos casos (179). En contraste, varios países que tenían niveles altos o intermedios de prevalencia de síntomas en la fase uno mostraron aumentos significativos en la prevalencia en la fase tres. Los ejemplos incluyen países latinoamericanos como Costa Rica, Panamá, México, Argentina y Chile, y países de Europa del Este como Ucrania y Rumanía.

Antes de la fase tres de ISAAC, había habido pocos estudios de cambios en la rinitis o la polinosis a lo largo del tiempo. En los centros del grupo de edad de 13 a 14 años en 56 países y en los centros del grupo de edad de 6 a 7 años en 37 países estudiados, se observó un ligero aumento mundial en la prevalencia de rinoconjuntivitis, pero las variaciones fueron grandes entre los centros y no hubo un patrón regional consistente (180).

En el momento de la fase tres de ISAAC, no estaba claro si la prevalencia del eccema estaba aumentando realmente en todo el mundo. En niños de 13 a 14 años de edad de 105 centros de 55 países y niños de 6 a 7 años de edad en 64 centros de 35 países, los cambios anuales de prevalencia en relación con la prevalencia promedio en las fases uno y tres fueron generalmente pequeños y difirieron en dirección según la edad de los participantes y la región mundial (181). Para los jóvenes de 13 a 14 años, la prevalencia de los síntomas del eccema disminuyó en algunos centros de prevalencia previamente altas del mundo desarrollado, como el Reino Unido y Nueva Zelanda, mientras que los centros con tasas de prevalencia previamente altas de los países en desarrollo siguieron aumentando. En los niños de 6 a 7 años de edad, la mayoría de los centros mostraron un aumento en los síntomas actuales del eccema. Patrones similares a estos estaban presentes para el eccema severo en ambas edades. Por lo tanto, la epidemia de eccema parece estar nivelándose o

disminuyendo en algunos países con tasas de prevalencia previamente altas. El panorama en otros lugares es mixto, y muchos países en desarrollo que antes eran de baja prevalencia experimentaban aumentos sustanciales, especialmente en el grupo de edad más joven.

En España el estudio ISAAC en su fase I fue seguida por 10 centros distintos, y participaron en ella 30.626 niños de 13-14 años y 19.010 de 6-7 años. Los cuestionarios dirigidos a los niños mayores fueron autocumplimentados; y los dirigidos a los niños pequeños fueron rellenados por los padres. España fue el país que más niños reclutó en ambos grupos de edad (182). La puesta en marcha de la fase II del estudio incluía cuestionario de síntomas y de factores de riesgo y tratamiento, examen de piel, prick-test cutáneo, prueba de provocación bronquial, determinación de inmunoglobulina E (IgE) total y específica, y medición de ácaros y endotoxinas bacterianas en polvo doméstico. Se repartió un cuestionario junto con la petición de autorización de los padres, al número de colegios necesario (clase de cuarto y quinto de primaria) para obtener alrededor de 1.000 autorizaciones para la realización de la prueba de prick-test en cada ciudad. La participación en el cuestionario, examen de piel, prick-test, extracción de sangre, provocación bronquial y recogida de polvo fueron, respectivamente, para cada ciudad: Cartagena, 58,9; 49,2; 43,2; 33,1; 39,9 y 29,8%; Almería, 42,5; 40,6; 39,1; 39,1 y 37,6%; Valencia, 43,7; 30,4; 23,5; 21,5; 20,1 y 17,1%; Madrid, 53,2; 40,8; 38,6; 27,6; 30,6 y 21,1%. En Cartagena, la participación fue significativamente mayor en los colegios concertados, así como en los del área no contaminada. La participación fue muy variable, aunque más baja que en los centros europeos en los que el estudio ha finalizado. La zona y el tipo de colegio influyen significativamente en la participación (183).

En la fase III se intentó averiguar cuál es la tendencia de la prevalencia del asma y de las enfermedades alérgicas en dos grupos de edad pediátrica o si las diferencias geográficas o por edades se mantienen en España (184). El objetivo era describir la prevalencia de los síntomas de asma en distintas áreas geográficas españolas mediante estudio transversal de prevalencia de asma, realizado en 2016-2019 a 19.943 adolescentes de 13-14 años y 17.215 escolares de 6-7 años de 6 áreas

geográficas españolas (Cartagena, Bilbao, Cantabria, La Coruña, Pamplona y Salamanca). El estudio GAN (Global Asthma Network) ha permitido actualizar la prevalencia de síntomas de asma en España, confirmando que el asma sigue siendo una patología común con una prevalencia de sibilancias recientes del 15,3% en los adolescentes y del 10,4% en los escolares, y que son superiores a las del ISAAC-III 2002-2003 (10,6% y 9,9%, respectivamente). Mientras en los adolescentes existe un incremento de la prevalencia, en los escolares parece estabilizarse. Aunque este aumento de la prevalencia de asma en los adolescentes pueda ser explicado por el menor número de centros, participantes y la ausencia de algunos centros. Los síntomas de asma se recogieron mediante un cuestionario escrito y videocuestionario según el protocolo Global Asthma Network (GAN). La prevalencia de sibilancias recientes (últimos 12 meses) fue del 15,3% a los 13-14 años y del 10,4% a los 6-7 años, con variaciones en los adolescentes, desde un 19% en Bilbao hasta un 10,2% en Cartagena; y en los escolares, desde un 11,7% en Cartagena hasta un 7% en Pamplona. Estas prevalencias fueron superiores a las del ISAAC (10,6% en adolescentes y 9,9% en los escolares). Un 21,3% de adolescentes y un 12,4% de los escolares refirieron asma alguna vez. La conclusión fue que existía una alta prevalencia de síntomas asmáticos con un incremento en los adolescentes y una estabilización en los escolares españoles con respecto al ISAAC. No se aprecian tan claramente variaciones geográficas en la prevalencia de asma, pero las áreas que tenían prevalencias elevadas mantienen cifras altas.

1.5. Datos no estructurados: un mundo por explorar, *Big data*

1.5.1. Importancia de la cantidad de datos producidos a nivel mundial

Big Data es uno de los temas actuales que está tomando relevancia en todas las industrias. Se considera que es uno de los temas más importantes ya que muchas empresas se están enfrentando al crecimiento del volumen, velocidad y variedad de datos. Además, existen áreas de oportunidad en la exploración de datos para la creación del valor.

Los seres humanos estamos creando y almacenando información constantemente y cada vez en más cantidades grandes (185). Esta contribución a la acumulación

masiva de datos la podemos encontrar en diversas industrias, las compañías mantienen grandes cantidades de datos transaccionales, reuniendo información acerca de sus clientes, proveedores u operaciones. De la misma manera sucede con el sector público. En muchos países se administran enormes bases de datos que contienen datos de censo de población, registros médicos, impuestos y si a todo esto le añadimos transacciones financieras realizadas en línea o por dispositivos móviles, análisis de redes sociales, ubicación geográfica mediante coordenadas, en otras palabras, todas aquellas actividades que la mayoría de nosotros realizamos varias veces al día con nuestros teléfonos móviles, estamos hablando de que se generan alrededor de varios quintillones de bytes diariamente en el mundo (186).

Pero no solamente somos los seres humanos quienes contribuimos a este crecimiento enorme de información, existe también la comunicación denominada máquina a máquina cuyo valor en la creación de grandes cantidades de datos también es muy importante: sensores digitales instalados en contenedores para determinar la ruta generada durante una entrega de algún paquete y que esta información sea enviada a las compañías de transporte, sensores en medidores eléctricos para determinar el consumo de energía a intervalos regulares para que sea enviada esta información a las compañías del sector energético, etc . Se estima que hay más de 30 millones de sensores interconectados en distintos sectores como automotriz, transporte, industrial, servicios, comercial y se espera que este número crezca en un 30% anualmente.

El crecimiento en el volumen de datos generados por diferentes sistemas y actividades cotidianas en la sociedad ha forjado la necesidad de modificar, optimizar y generar métodos y modelos de almacenamiento y tratamiento de datos que suplan las carencias que presentan las bases de datos y los sistemas de gestión de datos tradicionales. Respondiendo a esto aparece *Big Data*, término que incluye diferentes tecnologías asociadas a la administración de grandes volúmenes de datos provenientes de diferentes fuentes y que se generan con rapidez. A pesar de que el término *Big Data* se asocia principalmente con cantidades de datos exorbitantes, se debe dejar de lado esta percepción, pues *Big Data* no va dirigido solo a gran tamaño, sino que abarca tanto volumen como variedad de datos y velocidad de acceso y procesamiento. Es una referencia a aquellos sistemas de información que manejan

conjuntos de datos de gran volumen, de alta velocidad, de veracidad, de valor y de gran variedad de recursos, que demandan formas rentables e innovadoras de procesamiento de la información para mejorar la comprensión y la toma de decisiones. Es la solución al crecimiento exponencial de los datos, en el momento en que se hace difícil su administración con respecto al almacenamiento, procesamiento y acceso (187). Surge como una nueva era en la exploración y utilización de datos y desde la perspectiva empresarial no representa solo grandes volúmenes de datos, sino que se deben considerar los patrones extraídos a partir de los datos y que pueden generar procesos de innovación. En la actualidad se ha pasado de la transacción a la interacción, con el propósito de obtener el mejor provecho de la información que se genera minuto a minuto.

Con el auge del *Big Data* se ha dado cabida también a un nuevo concepto, *Data Science* o Ciencia de los Datos, que se usa de forma genérica para hacer referencia a la serie de técnicas necesarias para el tratamiento y manipulación de información masiva desde un enfoque estadístico e informático. Incluye también el surgimiento de un nuevo perfil profesional, el "*Data Scientist*", las personas capacitadas en este perfil deben saber de economía, de informática y de análisis e interpretación estadística.

Partiendo de la necesidad de almacenamiento y análisis de los datos se han desarrollado sistemas de archivos y otras herramientas (187). Las compañías que han sido pioneras en el uso de analíticas profundas sobre grandes bases de datos han sido las que operan sobre internet, como son los motores de búsqueda, los sitios de redes sociales y los sitios de comercio en línea. Sin embargo, el desarrollo de nuevos tipos de sensores remotos como telescopios, videocámaras, monitores de tráfico, máquinas de resonancia magnética, sensores químicos y biológicos y sensores de monitoreo ambiental, se han generado nuevos flujos de datos digitales. Así mismo, las personas a través de sus teléfonos móviles, ordenadores personales y otro tipo de dispositivos digitales generan grandes flujos de datos personales. Lo anterior deja ver que se presentan oportunidades incalculables para la formulación de investigación científica, acelera la innovación y puede ayudar a mejorar ámbitos que van desde la salud hasta el gobierno. También se abren nuevas oportunidades de negocio porque surgen mecanismos que permiten entender las dinámicas de

negocio en tiempo real, como el comportamiento de los consumidores, las actividades de vida nocturna, los mercados, entre otros (188).

Para la creación de soluciones en el manejo de grandes volúmenes de datos, se pueden encontrar cuatro fases donde se agrupan o clasifican las diferentes tecnologías de soporte, estas son: generación, adquisición, almacenamiento y análisis de datos. La primera fase, es la generación, como un proceso propio de diversas actividades de la sociedad, en estas se genera una cantidad inmensa de datos que, según su naturaleza, puede estar almacenada y estructurada o puede corresponder a datos sin ninguna estructura, pero con características de gran valor. En la segunda fase, se incluye la colección de todos estos datos generados en la vida diaria, la transmisión y pre-procesamiento de los mismos es de gran importancia, ya que muchos conjuntos de datos presentan redundancia o datos inútiles y si no se tratan pueden incrementar el espacio de almacenamiento innecesariamente y afectar los resultados de una fase de análisis. La fase de almacenamiento ha generado la necesidad de crear estudios y propuestas de nuevas estrategias que permitan afrontar los tipos de datos que no se pueden gestionar con un sistema de gestión de bases de datos relacionales. Surgen entonces, tecnologías de almacenamiento de datos masivos como almacenamiento con conexión directa y el almacenamiento en red. Finalmente, la fase de análisis debe atender a la necesidad de extraer rápidamente información desde los datos masivos para poder generar valor en las organizaciones y facilitar procesos de toma de decisiones, se requiere de tecnologías que faciliten incluso el análisis en tiempo real.

Por todos estos motivos se ha necesitado crear librerías para procesamiento de datos, bases de datos y modelos de programación junto con estrategias de procesamiento de grandes conjuntos de datos que puede ser aplicado a una gran variedad de tareas del mundo real (189). Además, han aparecido modelos de aprendizaje máquina y de inteligencia artificial (190).

La denominada Ciencia de Datos (*Data Science*) se ha vuelto muy popular hoy en día. Se trata de un campo multidisciplinar, con importantes aportaciones estadísticas e informáticas, dentro del que se incluirían disciplinas como minería de datos (*Data Mining*), aprendizaje automático (*Machine Learning*), aprendizaje

profundo (*Deep Learning*), modelado predictivo (*Predictive Modeling*), extracción de conocimiento (*Knowledge Discovery*) y también el aprendizaje estadístico (*Statistical Learning*) (191).

Podríamos definir la Ciencia de Datos como el conjunto de conocimientos y herramientas utilizados en las distintas etapas del análisis de datos. Otras definiciones podrían ser: el arte y la ciencia del análisis inteligente de los datos, también el conjunto de herramientas para entender y modelizar conjuntos (complejos) de datos, y por último el proceso de descubrir patrones y obtener conocimiento a partir de grandes conjuntos de datos. Aunque esta ciencia incluiría también la gestión (sin olvidarnos del proceso de obtención) y la manipulación de los datos (186,187). Podemos hacer una clasificación de las técnicas de Big data en técnicas estadísticas, métodos de optimización, minería de datos, técnicas de aprendizaje máquina, técnicas de clasificación y técnicas de análisis y regresión (192).

Minería de datos (Data Mining) intenta entender conjuntos de datos (en el sentido de encontrar sus patrones), requiere de una intervención humana activa (al igual que la inferencia estadística tradicional), pero utiliza entre otras las técnicas automáticas de aprendizaje máquina. Por tanto, podríamos pensar que es más parecido al aprendizaje estadístico. La minería de datos se puede definir como el proceso de extracción de conocimiento a partir de cúmulos de datos. Se suele utilizar el término minería de datos como sinónimo de descubrimiento de conocimiento, pero realmente no son sinónimos, la minería de datos es solo un paso en el proceso de descubrimiento de conocimiento. La minería de datos nace de la necesidad de conocer información útil a partir de las bases de datos. Con el crecimiento de los datos disponibles, la inteligencia de negocios tuvo que dar paso a la aplicación de la minería de datos en soluciones empresariales y comerciales, puesto que de esta manera se permite el descubrimiento automático o semiautomático de información relevante a partir de estos cúmulos de datos. En las ciencias y la ingeniería existe un amplio rango de problemas y dominios aplicación para la minería de datos. Se encuentran soluciones a partir de de minería de datos para problemas de los campos de comercio, salud, predicción, transporte, meteorología, entre otros (193).

El término *Machine Learning* (aprendizaje automático) se utiliza en el campo de la inteligencia artificial desde 1959 para hacer referencia, fundamentalmente, a algoritmos de predicción (inicialmente para reconocimiento de patrones). Es un conjunto de algoritmos principalmente dedicados a hacer predicciones y que son esencialmente automáticos minimizando la intervención humana. Es un área de investigación bastante reconocida en las ciencias de la computación y principalmente comprende el descubrimiento de modelos, patrones y regularidades en los datos (194). El aprendizaje máquina puede ser visto desde dos enfoques, los simbólicos y los estadísticos. Los primeros trabajan aprendizaje inductivo de descripciones simbólicas, mientras que los segundos se centran en los métodos de reconocimiento de patrones en la estadística. En los últimos años, el uso del aprendizaje máquina se ha extendido con rapidez, se ven aplicaciones en dominios como detección de fraudes, sistemas de recomendación, predicciones financieras y comercio entre otros. Los algoritmos de aprendizaje máquina se clasifican en supervisados y no supervisados (195).

Una de estas etapas (que están interrelacionadas) es la construcción de modelos a partir de los datos para aprender y predecir. Podríamos decir que el aprendizaje estadístico (AE) se encarga de este problema desde el punto de vista estadístico. En estadística se considera emplear modelos estocásticos (con componente aleatorio), para tratar de tener en cuenta la incertidumbre debida a que no se disponga de toda la información (sobre las variables que influyen en el fenómeno de interés). La inferencia estadística proporciona herramientas para ajustar este tipo de modelos a los datos observados (seleccionar un modelo adecuado, estimar sus parámetros y contrastar su validez). Los avances en informática han permitido el uso de modelos estadísticos más avanzados, principalmente métodos no paramétricos. Muchas de las herramientas que utilizan provienen del campo de la estadística y, en cualquier caso, la estadística (y por tanto las matemáticas) es la base de todos estos enfoques para analizar datos. Por este motivo se introdujo el término aprendizaje estadístico para hacer referencia a este tipo de herramientas (196).

Tradicionalmente no se preocupaba del origen de los datos, e incluso es habitual que se considere que un conjunto enorme de datos es equivalente a disponer de

toda la información. Por el contrario, en el caso del aprendizaje estadístico se trata de comprender, si es posible, el proceso subyacente del que provienen los datos y si estos son representativos de la población de interés (si tienen algún tipo de sesgo) (197). Dentro de los problemas que aborda el aprendizaje estadístico se suelen diferenciar dos grandes bloques: el aprendizaje no supervisado y el supervisado (198).

El aprendizaje no supervisado comprende los métodos exploratorios, es decir, aquellos en los que no hay una variable respuesta (al menos no de forma explícita). El principal objetivo de estos métodos es entender las relaciones entre los datos y su estructura, y pueden clasificarse en las siguientes categorías: análisis descriptivo, métodos de reducción de la dimensión (análisis de componentes principales, análisis factorial...), métodos de agrupación (conglomerados o clusters).

El aprendizaje supervisado engloba los métodos predictivos, en los que una de las variables está definida como variable respuesta. Su principal objetivo es la construcción de modelos que posteriormente se utilizarán, sobre todo, para hacer predicciones. Dependiendo del tipo de variable respuesta se diferencia entre: respuesta categórica (también se emplea la denominación de variable cualitativa, discreta o factor) y respuesta numérica (cuantitativa).

1.5.2. Tipos de datos

Los datos en sí mismos tienen limitada importancia, es necesario que nos den respuestas, en ello consiste, en esencia, el análisis e interpretación de los datos. El propósito del análisis es resumir las observaciones llevadas a cabo de forma tal que proporcionen respuesta a los interrogantes de la investigación. La interpretación, más que una operación distinta, es un aspecto especial del análisis. La relación entre análisis e interpretación y la forma específica que toman, tanto separada como conjuntamente, varían de un estudio a otro, dependiendo de los distintos esquemas o niveles de investigación y, fundamentalmente, del diseño propuesto (199).

Los datos, a partir de los cuales el investigador inicia el análisis, son diferentes según el nivel de elaboración realizado, el cual depende de la naturaleza del problema de

investigación y, consecuentemente, del tipo de investigación; también de las técnicas y procedimientos seguidos en la elaboración.

Se diferencian dos objetivos en el análisis de datos: información (en el sentido de inferencia) y predicción. Cada uno de estos objetivos da lugar a una cultura:

Modelización de datos: desarrollo de modelos (estocásticos) que permitan ajustar los datos y hacer inferencia. Es el trabajo habitual de los estadísticos académicos.

Modelización algorítmica (en el sentido de predictiva): esta cultura no está interesada en los mecanismos que generan los datos, solo en los algoritmos de predicción. Es el trabajo habitual de muchos estadísticos industriales y de muchos ingenieros informáticos.

De acuerdo a estas consideraciones, los datos que se utilizan en el análisis pueden ser: datos estructurados (datos cuantificados, datos no cuantificados), datos no estructurados (200).

1) Datos estructurados

Los que pueden estar ordenados o contenidos en una tabla.

A) Análisis de los datos cuantificados.

Algunos tipos de estudios, por su naturaleza, aportan datos elaborados, es decir, cuantificados. El tratamiento estadístico de los datos permite un análisis adecuado que puede tener diversos alcances, los cuales dependen de los objetivos de la investigación y de las hipótesis formuladas.

El análisis puede estar orientado a:

Determinar lo que es típico en el grupo estudiado (se utiliza algunas de las medidas de tendencia central, según el caso).

Indicar si existen variaciones entre los sujetos del grupo, señalando de qué tipo y magnitud (se utiliza alguna de las medidas de variabilidad; cada una proporciona datos sobre un aspecto diferente).

Mostrar la forma cómo están distribuidos los individuos con respecto a la variable que se mide (se utiliza el desarrollo de una curva de distribución).

Mostrar la relación existente entre dos o más variables (se aplica el coeficiente de variabilidad).

Describir las diferencias existentes comparando dos grupos de individuos.

B) Análisis de los datos no cuantificados.

No todos los aspectos del material recogido pueden ser categorizados y, consecuentemente, cuantificados, debido, en algunos casos, a la falta de precisión en la definición de las categorías, lo que dificulta el análisis de los resultados. Por este motivo, se recomienda considerar que cada categoría propuesta comprenda un amplio margen de criterios para las respuestas.

2) Análisis de los datos no estructurados.

El material no estructurado es el que proviene, por ejemplo, de observaciones o entrevistas no estructuradas, en las cuales se recoge mucho material, a veces valioso, pero sin ninguna pauta que permita alguna forma de organización y menos de clasificación. En algunos casos, los estudios de nivel exploratorio, que no se inician con hipótesis, cubren aspectos diversos, los cuales conducen al acopio de datos en cantidad excesiva y no estructurada. El problema que plantea este tipo de datos es doble: primero porque se necesita determinar qué aspectos del material requieren ser categorizados, y segundo, saber qué principios de clasificación pueden utilizarse. Se propone, como soluciones al problema expuesto: elaborar, en primer término, hipótesis de trabajo que permitan establecer principios de clasificación y, en segundo lugar, utilizar algunos procedimientos que puedan ayudar en el análisis, tales como estudiar el material correspondiente a un grupo que contrasta con el que se está investigando, con el objeto de obtener elementos que sugieran ideas sobre las diferencias significativas entre ambos grupos, respecto de la característica que se analiza. Otro procedimiento consiste en formar grupos con los casos motivo de estudio sobre características comunes; después se analizan para ver si aquellos que tienen características semejantes han pasado por experiencias parecidas.

También, pueden formarse grupos sobre la base de aquellos que han tenido experiencias similares, y ver en qué medida, son semejantes respecto de las características comunes que presentan.

El objetivo de la interpretación es buscar un significado más amplio a las respuestas mediante su unión con otros conocimientos disponibles. Ambos propósitos, por supuesto, presiden la totalidad del proceso de investigación, todas las fases

precedentes han sido tomadas y ordenadas para hacer posible la realización de estos dos últimos momentos.

Debe manejarse con prudencia los valores obtenidos con el análisis estadístico de los datos, pues no siempre la significación estadística de los resultados garantiza que éstos sean realmente importantes. Cuando los resultados no confirman las hipótesis, el investigador debe, sin embargo, aceptarlos como tales, puesto que en esa condición tendrán su propio significado y valor. Debe aplicarse por lo común alguna prueba de significación a los datos, para verificar la hipótesis. En la medida de lo posible, la tabla o tablas que presentan los resultados de un estudio deben explicarse por sí mismas y no requerir una extensa lectura del texto para ser interpretadas. En cambio, el material del texto debe señalar los aspectos importantes de los datos y atraer la atención hacia la pertinencia de los resultados.

En sanidad el análisis de datos no estructurados es de gran valor. Los datos estructurados responden a un marco de conocimiento que ya tenemos. El análisis de datos no estructurados nos permite conocer de una manera no invasiva la calidad percibida de pacientes y ciudadanos tanto sobre el sistema como sobre su propia salud. Esto resulta clave cuando queremos aumentar el valor que el sistema sanitario produce para las personas y para la población. La situación en España es de un aparente desinterés. El sistema de seguridad social tiene más de 60 millones de datos. El valor de estos datos sanitarios es inmenso. Legislar y gestionar este capital, evitando una apropiación indebida y facilitando la generación de valor para los ciudadanos debería ser una prioridad. Si no, puede llegarse a una situación en la que las multinacionales sanitarias nos cobren mañana por el conocimiento sanitario que hoy nos han extraído a cambio de nada. Es preciso abrir un debate sobre la propiedad del dato sanitario, incluso sobre la legalidad de que el dato sanitario pueda ser propiedad de alguien (201).

1.5.3. Forma de analizar los datos

En general, cualquiera que haya intentado registrar y posteriormente analizar en forma más o menos avanzada sus datos, se encuentra con un problema: como describir la información contenida en un conjunto de datos mediante algún conjunto menor de variables que sean manipulables y signifiquen algo para el investigador.

La idea subyacente en el método que analizaremos es que si una variable es función de otras, contiene información redundante. Por lo tanto, si muchas de las variables observadas están fuertemente correlacionadas, será posible sustituirlas por menos variables sin gran pérdida de la información. El problema práctico sería el típico de una hoja de Excel en la que tenemos registrados gran cantidad de pacientes y gran cantidad de variables de cada paciente: cómo hacer que los datos nos muestren algún patrón que los relacione y a su vez elimine los datos de alguna manera redundantes (202).

Las técnicas estadísticas multivariadas son aquellas que analizan múltiples características, medidas en un mismo individuo. El análisis multivariado es una técnica matemática que permite el estudio simultáneo de las relaciones entre más de dos variables y de las unidades de análisis en un período de tiempo determinado. Constituyen una generalización de las técnicas univariadas y bivariadas, donde todas las variables deben ser aleatorias y estar interrelacionadas de tal forma que no tenga sentido interpretar de forma aislada sus diferentes efectos.

Las estadísticas univariadas (una sola variable), son empleadas con frecuencia en los estudios sociales. En estos casos es suficiente conocer la media y la varianza del objeto en estudio para determinar la distribución de la variable y de este modo inferir los resultados de la investigación mediante las pruebas de hipótesis y la conformación de intervalos de confianza. El avance de la tecnología informática ha hecho posible que en las investigaciones se apliquen modelos analíticos complejos para estudiar los fenómenos sociales, dando lugar al uso frecuente de las técnicas multivariadas. Esto facilitó el estudio simultáneo de las variables y una mejor aproximación al conocimiento de la realidad social. Mediante el análisis multivariado se pueden analizar más de dos variables en forma simultánea, generándose por cada variable una media y varianza, dando lugar a la formación de la matriz de covarianzas. La formulación de un modelo analítico, mediante el cual se hacen explícitas las relaciones entre más de dos variables, y el establecimiento de algunos supuestos previos, facilitan la interpretación posterior de los resultados (203). Las técnicas multivariadas son herramientas que permiten al investigador extraer abundante información de los datos disponibles. Las mismas son complejas y requieren para su utilización de un conocimiento profundo de sus fundamentos.

Los paquetes estadísticos han permitido una amplia aplicación de las mismas, aunque algunas veces por insuficiente conocimiento y respaldo teórico, su empleo no ha sido el adecuado (204).

La técnica multivariada a utilizar se determina a partir del conocimiento de la forma en que se relacionan las variables inherentes al fenómeno social en estudio. El análisis presenta restricciones según el tipo de variables existiendo una técnica específica de acuerdo al tipo de variables investigadas. Por ejemplo el análisis factorial y de conglomerados se aplica generalmente cuando las variables son cuantitativas, mientras que el análisis discriminante exige que la variable dependiente sea cualitativa y las independientes cuantitativas o dicotómicas. En cuanto a las unidades de análisis el análisis es de corte transversal (en un determinado período de tiempo) y permite el estudio simultáneo de todas las unidades de análisis.

Los modelos de análisis multivariado son empleados con más frecuencia en los estudios observacionales acerca de la etiología y el pronóstico de una enfermedad, y en ocasiones en los estudios de diagnóstico o de intervención. El análisis multivariado está presente en los estudios de causalidad, cuya estrategia debe partir de un buen diseño, donde se involucren como variable dependiente el desenlace y como covariables los supuestos factores causales, donde son importantes los posibles factores confusores que deben incluirse en el análisis y se toman precauciones para evitar posibles sesgos. En los estudios observacionales, donde no es posible hacer una asignación aleatoria de los individuos a los diferentes grupos de estudio, es muy útil el análisis multivariado, ya que permite el ajuste de las características basales de los grupos para lograr una aproximación estadística razonable a la estimación del efecto de la intervención, una mayor economía en el análisis de los datos, el control de las variables confusoras y obtener una mayor consistencia en la inferencia estadística (204).

1.5.3.1. Clasificación de las técnicas estadísticas multivariadas.

Para saber qué técnica se debe aplicar, es necesario conocer cómo están medidas las variables que se analizan en el estudio. También es preciso conocer el tipo de relación existente entre las variables. La dependencia o no entre variables se

determina por el conocimiento *a priori* del investigador. Como consecuencia de estos aspectos, los principales métodos de análisis multivariado empleados en la inferencia causal, se pueden agrupar según su finalidad (205).

1.5.3.1.1. **Métodos explicativos o de dependencia:** son empleados para analizar las relaciones de dependencia que permiten explicar o predecir determinados eventos (206).

Según el tipo de variables dependientes pueden ser:

---Con variable dependiente cuantitativa:

Para explicar la variabilidad individual se puede emplear la regresión lineal múltiple, en el caso de una sola variable dependiente, y la regresión lineal multivariada en el caso de más de una variable dependiente. Para explicar la variabilidad grupal (los grupos condicionan la variable categórica) se emplea el análisis de varianza o de covarianza (ANOVA o ANCOVA) si es una sola variable o un análisis multivariado de varianza o de covarianza (MANOVA o MANCOVA) si se trata de más de una variable dependiente.

En los estudios observacionales, en los que no es posible asignar experimentalmente valores de una variable predictiva, al investigador le interesa conocer los efectos de múltiples variables predictivas sobre un resultado y para ello puede usar los modelos de regresión que sintetizan la información sobre variables predictivas individuales con el uso de todo el conjunto de datos. Los coeficientes estimados con el modelo son parámetros estadísticos descriptivos muy potentes que permiten obtener un significado de los datos. Los métodos de regresión múltiple se emplean según el tipo de variable dependiente que se analiza, así se tienen, entre los más empleados:

Regresión lineal: cuando la variable dependiente es continua. Los cambios de la covariable (aumento o disminución) influyen en el valor esperado de la variable dependiente de una manera lineal. Cada coeficiente de regresión representa el cambio que se espera en la variable de resultado por cada unidad de cambio en la covariable.

Regresión logística: cuando la variable dependiente es categórica o cualitativa. En su forma binaria es muy empleada en la identificación de factores de riesgo o pronósticos cuando la variable dependiente es dicotómica, por ejemplo, vivo o fallecido o enfermo o no enfermo, y se pretende encontrar la probabilidad de que ocurra el desenlace en estudio con la influencia o no de otras variables.

También puede ser multinomial cuando permite tratar situaciones en las que la variable respuesta tienen más de dos alternativas posibles, sean ordinales o no.

Regresión de Cox (modelo de riesgos proporcionales): empleada en los estudios de supervivencia, donde la variable dependiente está en función del tiempo transcurrido hasta un determinado evento y se quiere determinar simultáneamente el efecto de una serie de factores que pueden alterar dicho tiempo. Ese efecto se presenta como la razón de riesgo (HR de Hazard Ratio) y expresa la magnitud en la que una variable aumenta o disminuye el riesgo de ocurrencia de un desenlace en el tiempo, por lo que supera a la técnica bivariada de Kaplan Meier.

Regresión binomial: empleada cuando la variable de respuesta es dicotómica, es decir, el resultado solo puede adoptar uno de dos valores posibles como enfermar o no enfermar, morir o sobrevivir. En estos casos modela el logaritmo de una proporción, que bien podría ser la incidencia o la prevalencia de una enfermedad.

Regresión de Poisson: empleada cuando la variable dependiente se ajusta bien a una distribución de Poisson (distribución que representa la probabilidad de que un determinado número de eventos ocurra durante un periodo de tiempo en un espacio o población especificada), para cualquier combinación de valores de la covariable. Una variable Poisson se presentan como la tasa de un evento de interés, por ejemplo, el número de fallecidos por determinada causa en una población específica durante el periodo de observación. Esta técnica modeliza bien situaciones de conteo en una unidad de tiempo.

Regresión discontinua: apropiada en diseños cuasi-experimentales en los que se investiga el efecto de las intervenciones asignadas por encima o por debajo de un

valor de corte o umbral. Mediante la comparación de las observaciones en un entorno del valor umbral, es posible estimar el valor del efecto de tratamiento promedio en los casos que la aleatorización es imposible.

---Con variable dependiente cualitativa:

El análisis discriminante para predecir si un individuo pertenece a un grupo cuando se cumple el supuesto de normalidad. El análisis discriminante es otra técnica del análisis multivariado que permite clasificar las unidades de análisis en grupos definidos *a priori* y analizar las causas que han dado lugar a la formación de estos grupos. Los grupos se conforman a partir de un conjunto de variables seleccionadas (variables independientes), las cuales además de explicar la formación de los grupos pueden ordenarse según su mayor poder discriminatorio. De este modo las unidades de análisis son reclasificadas, generándose una clasificación distinta a la realizada *a priori*, donde el aporte de las variables independientes (explicativas) a la formación de los grupos determina la formación de los mismos (207).

1.5.3.1.2. **Métodos descriptivos o de interdependencia**: son empleados para verificar las relaciones de interdependencia y describir la estructura de los datos según se quieran identificar (208). Se dividen en:

TECNICAS DE REDUCCION

---Variables similares con variables cuantitativas: se emplean el análisis de componentes principales (ACP) y el análisis factorial común (AFC). Estas técnicas permiten reducir la dimensionalidad de los datos, con lo que se pueden descartar atributos irrelevantes para la toma de decisiones, que van en contra de la eficiencia de los algoritmos propios de las técnicas empleadas *a posteriori*. Las nuevas características encontradas son combinación lineal de las anteriores y se les llama componentes si se forman por variables observadas y factores si se forman por variables no observadas (209). Las dos técnicas forman parte de un conjunto de técnicas multivariadas denominadas de “interdependencia”, ya que analizan la relación mutua entre un conjunto de variables. Su finalidad principal, por lo tanto, no es el análisis de relaciones causales, sino la agrupación de variables, en función

de la variabilidad que cada variable comparte con otras variables (varianza o covarianza). Se busca la síntesis de la información proporcionada por cierto grupo de variables observadas, en un número inferior de variables no observadas (factores comunes o componentes principales, depende de la variedad analítica que se realice), con la menor pérdida de información posible. Dichas variables no observadas (o latentes), darían cuenta de conceptos no observables directamente, que engloban lo que tienen en común las variables observadas.

El análisis de componentes principales (ACP) se caracteriza por analizar la varianza total del conjunto de variables observadas. De ellas trata de determinar las dimensiones básicas (o “componentes”) que las definen. Se considera la varianza total de la serie de variables observadas. El propósito es maximizar la proporción total de la varianza explicada. La idea central del análisis de componentes principales (ACP) es reducir la dimensionalidad de un conjunto de datos consistente en un número elevado de variables interrelacionadas. Se trata de mantener de la mejor manera posible la variación contenida en los datos. Esto se logra transformando el conjunto original en un nuevo conjunto de datos, los componentes principales, que son no correlacionados y que están ordenados de tal manera que los primeros pocos retengan la mayor parte de la variación presente en todas las variables originales. Dentro de los métodos de análisis de más de una variable, el análisis de componentes principales es una forma rápida y sencilla de encontrar un patrón de relación entre los datos y a su vez de reducir gran cantidad de variables. Sus fundamentos modernos se encuentran en el álgebra lineal y su gran difusión en los últimos años se debe al poder de cómputo actual.

En el análisis de factor común (AFC) el estudio de las interrelaciones entre las variables se restringe, en cambio, a la varianza común (o covarianza), es decir, a la búsqueda de un número reducido de “factores” que expresen lo que es común al conjunto de variables observadas. Está orientado al análisis de la covarianza. El análisis factorial es una técnica de análisis multivariado que permite obtener a partir de un conjunto de variables cualitativas un grupo menor de nuevas variables denominadas factores, los mismos que estarían explicando la variación conjunta

dependencia mutua entre dichas variables. Estos factores denominados variables latentes se caracterizan por no estar correlacionados entre sí. Con esta reducción se hace más sencillo el análisis de los resultados. Los tipos más frecuentes del análisis factorial son: el análisis factorial exploratorio y el análisis factorial confirmatorio. El primero se utiliza cuando el investigador requiere clasificar a las variables en dimensiones excluyentes (factores).

Las técnicas antes mencionadas pueden ser de amplio uso en la práctica, ya que en la clínica es frecuente emplear, como variables, múltiples determinaciones de laboratorio entre otras variables cuantitativas con la cual se pueden hacer investigaciones y análisis multivariados en dependencia de lo que se desea probar. En gran parte de los problemas científicos en general y en particular en biología y medicina, las variables son multidimensionales (analizan varias variables). La gran utilidad de la técnicas multivariantes reside en el análisis de datos de elevada dimensionalidad, donde la graficación de las variables es dificultosa si no imposible (210).

---Con variables cualitativas se emplean el análisis de correspondencia y el modelo logístico lineal para tablas de contingencia con más de dos entradas.

TECNICAS DE AGRUPAMIENTO

---Con variables de cualquier tipo se emplean el análisis de conglomerados (*cluster*).

Es una técnica de agrupamiento que clasifica las unidades de la matriz de datos en conglomerados (grupos internamente homogéneos y heterogéneos entre ellos) con el empleo de diferentes índices de similaridad o de distancia.

En la matriz de correlaciones se pueden observar las correlaciones parciales, es decir, a nivel bivariado. Como buscamos colinealidad, esperamos que los valores fuera de la diagonal (los que no son la correlación de la variable con sí misma) sean mayores a 0,3 (baja colinealidad), más tendientes hacia 0,5 (colinealidad media), y óptimamente igual o mayor a 0,7 (colinealidad alta). Junto con la matriz de correlaciones, generamos el determinante de esta: un determinante bajo, es decir, cercano a 0, indica alta multicolinealidad entre las variables. No debe ser, sin

embargo, igual a cero (matriz no singular), pues esto indicaría que algunas de las variables son linealmente dependientes y no se podrían realizar ciertos cálculos necesarios para los procedimientos multivariados. En este caso observamos que es muy cercano a 0, lo que sugiere alto nivel de colinealidad en el conjunto de variables involucradas en la matriz.

Estas clasificaciones de las unidades de la matriz de datos en conglomerados pueden ser jerárquicas (ascendente o descendente) en los que no se establece *a priori* el número de grupos, o no jerárquicas en las que se parte de un número dado de grupos. El método no jerárquico o repartición comienza con una solución inicial, un número de grupos fijado de antemano y agrupa los objetos para obtener los grupos. Los métodos jerárquicos se subdividen a su vez en aglomerativos y divisivos: los métodos jerárquicos aglomerativos comienzan con tantos conglomerados como objetos tengamos que clasificar y en cada paso se recalculan las distancias entre los grupos existentes y se unen los dos grupos más similares. El algoritmo acaba con un conglomerado conteniendo todos los elementos. Los métodos jerárquicos divisivos comienzan con un conglomerado que engloba a todos los elementos y en cada paso se divide el grupo más heterogéneo. El algoritmo acaba con tantos conglomerados (de un elemento cada uno) como objetos se hayan clasificado.

En resumen, según el tipo exploratorio podemos clasificar los conglomerados:

Jerárquico aglomerativo:

- Análisis *cluster* de casos: se agrupan individuos.
- Análisis *cluster* de variables: se agrupan variables.

Algoritmo de optimización por k-medias: No genera un dendrograma, solo identifica los *clusters*.

Un *cluster*/conglomerado no es otra cosa que la agrupación de individuos en base a las semejanzas que tengan las medidas de sus observaciones. Es una técnica multivariante cuya idea básica es clasificar objetos formando grupos/conglomerados (*cluster*) que sean lo más homogéneos posible dentro de sí mismos y heterogéneos entre sí. Surge ante la necesidad de diseñar una estrategia que permita definir grupos de objetos homogéneos. Este agrupamiento se basa en la idea de distancia o similitud entre las observaciones y la obtención de dichos

conglomerados depende del criterio o distancia considerados y de lo que consideremos como similar. Resumiendo, el objetivo del análisis clúster es obtener grupos de objetos de forma que, por un lado, los objetos pertenecientes a un mismo grupo sean muy semejantes entre sí y, por el otro, los objetos pertenecientes a grupos diferentes tengan un comportamiento distinto con respecto a las variables analizadas (210).

Es una técnica exploratoria puesto que la mayor parte de las veces no utiliza ningún tipo de modelo estadístico para llevar a cabo el proceso de clasificación. Es una herramienta de exploración de datos que se complementa con técnicas de visualización de los mismos. Fundamentalmente es una tarea de clasificación. La clasificación es uno de los objetivos fundamentales de la ciencia y en la medida en que el análisis clúster nos proporciona los medios técnicos para realizarla, se nos hará imprescindible en cualquier investigación. Por ejemplo, clasificar grupos de consumidores respecto a sus preferencias en nuevos productos, clasificar las entidades bancarias donde sería más rentable invertir, clasificar las estrellas del cosmos en función de su luminosidad, identificar si hay grupos de municipios en una determinada comunidad con una tendencia similar en el consumo de agua con el fin de identificar buenas prácticas para la sostenibilidad y zonas problemáticas por alto consumo. Su aplicabilidad es evidente. Las técnicas de análisis *cluster* han sido tradicionalmente utilizadas en muchas disciplinas, por ejemplo, astronomía, comercio (segmentación de mercados, investigación de mercados), psicología, biología (taxonomía), ciencias ambientales (por ejemplo tipologías según la calidad de las aguas), sociología, economía, ingeniería. Como se puede comprender fácilmente el análisis clúster tiene una extraordinaria importancia en la investigación científica, en cualquier rama del saber.

Los inconvenientes del análisis de conglomerados es que es un análisis descriptivo, teórico y no inferencial, y habitualmente se utiliza como una técnica exploratoria que no ofrece soluciones únicas, las soluciones dependen de las variables consideradas y del método de análisis utilizado. Conviene estar siempre alerta ante el peligro de obtener, como resultado del análisis, no una clasificación de los datos sino una disección de los mismos en distintos grupos. El conocimiento que el

analista tenga acerca del problema decidirá que grupos obtenidos son significativos y cuáles no (208,210).

Proceso que se debe seguir en un análisis de conglomerados:

Paso 1: Selección de variables. La clasificación dependerá de las variables elegidas. Introducir variables irrelevantes aumenta la posibilidad de errores. Hay que utilizar algún criterio de selección: seleccionar sólo aquellas variables que caracterizan los objetos que se van agrupando, y referentes a los objetivos del análisis clúster que se va a realizar. Si el número de variables es muy grande se puede realizar previamente un análisis de componentes principales y resumir el conjunto de variables. A veces, en vez de usar todas las variables que tengamos en nuestro análisis, es recomendable que se reduzca su número. Esto se puede realizar, por ejemplo, ejecutando un análisis de componentes principales.

Paso 2: Detección de valores atípicos. El análisis *cluster* es muy sensible a la presencia de objetos muy diferentes del resto (valores atípicos).

Paso 3: Seleccionar la forma de medir la distancia entre objetos dependiendo de si los datos son cuantitativos o cualitativos. Datos métricos: medidas de correlación y medidas de distancia. Datos no métricos: medidas de asociación.

Paso 4: Estandarización de los datos (decidir si se trabaja con los datos según se miden o estandarizados). El orden de las similitudes puede cambiar bastante con sólo un cambio de escala de una de las variables por lo que sólo se realizará una tipificación cuando resulte necesario.

Paso 5: Obtención de los *clusters* y valoración de la clasificación realizada.

1.6. Encuesta Nacional de Salud

1.6.1. Generalidades

La Encuesta Nacional de Salud (ENS) es un estudio de la percepción del estado de salud, sus factores determinantes y la atención recibida, por parte de la población, incluida en el Plan Estadístico Nacional, realizado por el Instituto Nacional de Estadística (INE), por encargo del Ministerio de Sanidad y Consumo en virtud del Acuerdo Marco de colaboración entre ambas instituciones hasta la fecha, y con una

periodicidad que empezó siendo bianual y que después se ha transformado en cuatrienal. El trabajo de campo es llevado a cabo por el Centro de Investigaciones Sociológicas (CIS). El objetivo es disponer de la información necesaria para la más adecuada instauración de actuaciones en materia sanitaria (211, 212, 213).

Los aspectos fundamentales de estudio se pueden agrupar en cuatro:

1. Morbilidad percibida; que entre otros incluyen el estado general de salud, los accidentes, el grado de funcionalidad, las dolencias, etc.
2. Uso de los servicios de salud que contempla la consulta médica, los servicios de urgencias, las vacunaciones, la espera, consumo de medicamentos, etc.
3. Hábitos de vida como el consumo de tabaco y alcohol, las horas de sueño, la actividad física y los hábitos nutricionales, entre otros.
4. Aspectos específicos de la encuesta de menores, como el tipo de lactancia, la vacunación o el tiempo dedicado a la televisión.

Frente a los indicadores objetivos de salud que proporciona el estudio de la mortalidad, la percepción de la salud, que se obtiene a través de las impresiones del individuo, proporciona un elemento indispensable en la medición de la salud. La percepción de cada individuo refleja la contribución de distintos factores, sociales, ambientales y de estilos de vida que van más allá del sistema sanitario. El marco de las estrategias de salud que se desarrollan en nuestro país, cuyo objetivo es mejorar la salud de la población, precisa de este tipo de información subjetiva como otro elemento primordial para la planificación y adopción de medidas de Salud Pública. Es un estudio sobre el estado de salud y los factores determinantes del mismo desde la perspectiva de los ciudadanos, es decir, ofrece la visión que las personas tienen de su estado de salud y de la asistencia de atención primaria y especializada. Por tanto, permite completar el sistema de información sanitaria aportando datos de las personas que no acuden al sistema sanitario. Proporciona información sobre aspectos como la morbilidad percibida, los hábitos de vida, las conductas relacionadas con los factores de riesgo, la utilización de los servicios sanitarios y prácticas preventivas.

La ENS es una operación estadística incluida en el Plan Estadístico Nacional y que ha venido siendo realizada por el Ministerio de Sanidad y Consumo hasta el año 2001. En virtud del Acuerdo Marco de colaboración entre el Ministerio de Sanidad y

Consumo y el Instituto Nacional de Estadística para la realización de esta Encuesta, firmado el 11 de marzo de 2002, el INE se ha encargado en lo sucesivo de la realización del proyecto técnico, del diseño muestral, la preparación de los cuestionarios, la formación de entrevistadores, la recogida de los datos y el tratamiento informático hasta llegar a los ficheros finales de la encuesta.

Las Encuestas Nacionales de Salud de España (ENS) son definidas por el propio Ministerio de sanidad y Consumo y por el Instituto Nacional de Estadística como un “material para estudios de investigación del estado de salud de los españoles, así como un instrumento para la toma de decisión en las políticas sanitarias”. Se trata de macroencuestas realizadas en el ámbito geográfico nacional a personas residentes en viviendas familiares principales, quedando por tanto excluido la población residente en instituciones.

La Encuesta Nacional de Salud (ENS) se realizó por primera vez en el año 1987, continuando luego con carácter bianual, en 1993, 1995, 1997, 2001, 2003, 2006 y cuatrienal en 2011, 2017.

Las ENS son parte importante del Plan Estadístico Nacional y están integradas en el Sistema de Información del Sistema Nacional de Salud, ya que son un medio clave para el análisis de la salud de los ciudadanos y los comportamientos relacionados con la misma, constituyendo la base fundamental para la toma de decisiones en materia de salud pública.

El núcleo o contenido principal de las ENS se ha mantenido prácticamente sin cambios desde 1987, con el lógico objetivo de que el material pueda servir para elaborar estudios de evolución o tendencias de las variables reseñadas. Posteriormente, según se iban diseñando nuevas ediciones, las ENS han incorporado nuevos bloques temáticos y sus correspondientes preguntas, con el fin de actualizar los contenidos a las necesidades del momento y la evolución de las ciencias de la salud.

Desde el año 2003, por las necesarias exigencias de rigor estadístico, las ENS realizan la selección de la muestra de individuos por método aleatorio a partir del Censo de Población del Instituto Nacional de Estadística, siendo así el universo de estudio un conjunto claramente definido. Sin embargo, desde su primera edición

hasta el 2001 inclusive, la selección de las personas entrevistadas se realizó por técnicas no probabilísticas (cuotas sexos, cuotas edad).

Desde 1995 la información de las ENS se recoge a lo largo de los cuatro trimestres, para evitar la “estacionalidad” en los resultados; sistema no empleado en las ediciones de 1987 y 1993, no utilizadas en este trabajo.

El tipo de muestreo utilizado es trietápico estratificado. Las unidades de primera etapa son las secciones censales. Las unidades de segunda etapa son las viviendas familiares principales. Dentro de las viviendas se investigan todos los hogares que tienen su residencia habitual en ellas. Dentro de cada hogar se selecciona a una persona adulta (de 15 y más años) para cumplimentar el cuestionario individual. En el caso de que hubiera menores (de 0 a 14 años) se selecciona además un menor para cumplimentar el cuestionario de menores.

La encuesta consta de tres cuestionarios: Cuestionario de Hogar, Cuestionario de Adultos y Cuestionario de Menores. El método de recogida de información es el de entrevista personal asistida por ordenador (CAPI), directa en el caso de los adultos, y a la madre, padre o tutor en el caso de los menores de 15 años. El cuestionario consta de cuatro grandes bloques, sociodemográfico, estado de salud, asistencia sanitaria y determinantes de la salud.

1.6.2. Objetivos de las Encuestas Nacionales de Salud

La Encuesta Nacional de Salud tiene como objetivo general proporcionar la información necesaria sobre la salud de la población para poder planificar y evaluar las actuaciones en materia sanitaria, y proporcionar la adecuada atención a los usuarios de los servicios de salud.

Objetivos específicos:

1. Proporcionar información sobre la valoración del estado de salud general e identificar los principales problemas que sienten los ciudadanos (enfermedades crónicas, dolencias, limitaciones de la actividad, ayudas, accidentes).
2. Conocer el grado de utilización de los servicios de salud y sus características.
3. Conocer el uso de determinadas prácticas preventivas.

4. Conocer la frecuencia y distribución de hábitos de vida que suponen riesgo para la salud.
5. Conocer la percepción que los individuos tienen de su propia salud y la incidencia de algunas enfermedades y problemas de salud (problemas de salud que se pregunta al encuestado si han sido diagnosticados por el médico)
6. Relacionar la información anterior con las características sociodemográficas de la población.

Desde la ENS de 2006 se añaden al cuestionario preguntas referentes a determinantes de salud que no estaban contemplados en años anteriores; en concreto el medio ambiente físico y el medio ambiente social. También se incluyeron en la ENS de 2006 un nuevo campo, la salud mental, y se potencia la llamada “perspectiva de género” para detectar con mayor nitidez desigualdades de salud entre hombres y mujeres. También es necesario mencionar que desde la edición de 2006 se han reformulado algunas preguntas para iniciar la adaptación del formato de la encuesta a las necesidades del proyecto de la Encuesta Europea de Salud.

1.6.3. Contenido de las Encuestas Nacionales de Salud

El contenido de estas encuestas trata una serie de áreas de información comunes, entre las que destacan el estado de salud, los hábitos o estilos de vida, las prácticas preventivas, la utilización de los servicios sanitarios y en las últimas ediciones el medio ambiente.

-Estado de salud. Las ENS contienen en este sentido preguntas dirigidas a analizar los aspectos que a continuación se citan.

- Valoración subjetiva del propio estado de salud del encuestado
- Limitación de la actividad por enfermedad subjetivamente padecida
- Naturaleza de los síntomas o dolores que limitan la actividad
- Consumo de medicamentos
- Patologías diagnosticadas por un médico: hipertensión, colesterol elevado, diabetes, asma, bronquitis crónica, úlcera de estómago, enfermedades del corazón, alergia, depresión, otras enfermedades mentales, jaquecas o

migrañas, mala circulación, artrosis o reuma, osteoporosis, hernias, problemas menopáusicos distintos de la osteoporosis y problemas de próstata.

- Accidentabilidad. Utilización de servicios sanitarios y causa de la misma
- Visitas al odontólogo
- Utilización de servicios de urgencia y hospitalización
- Hábitos o estilos de vida
- Tabaquismo
- Consumo de bebidas alcohólicas
- Actividad y ejercicio físico
- Tipo de desayuno
- Frecuencia de consumo de los distintos grupos de alimentos
- Prácticas preventivas (vacunación gripe, visitas ginecólogo, citologías, mamografías, etc).
- Características auditivas y visuales
- Autonomía funcional en los mayores de 65 años

Como es lógico las ENS también relacionan los datos de los campos de información anteriormente mencionados con características sociales y demográficas de la población. La utilización de las ENS como base para la realización de estudios, o para la toma de decisiones en salud pública ofrece evidentes ventajas. En primer lugar, se trata de un conjunto de encuestas de metodología validada por el Instituto Nacional de Estadística, y avaladas por la experiencia desde 1987. En segundo lugar, proporcionan información sobre toda la población española, con independencia de que utilice o no los servicios sanitarios. Por otra parte, en la actualidad las ENS se consideran investigaciones polivalentes para la recogida de datos de salud y ofrecen un enfoque múltiple del fenómeno, ya que recogen tanto datos como la percepción del estado de salud, la existencia o no existencia de diagnóstico médico para varias patologías, los comportamientos o hábitos relacionados con la salud, la utilización de los servicios sanitarios, los aspectos ambientales del entorno y la calidad de vida.

1.6.4. Características fundamentales objeto de estudio

1.6.4.1. Morbilidad percibida por la población

- Estado general de salud, enfermedades crónicas, dolencias, molestias, limitaciones, tal como son vividas y expresadas por los ciudadanos. Se trata de medir la incidencia de cada una de esas carencias o limitaciones de salud y su repercusión en la vida de las personas.

- Asimismo se trata de medir la repercusión y trascendencia que tienen los accidentes de cualquier tipo que tiene la población encuestada: donde tienen lugar, efectos producidos y tipo de consulta médica ocasionada.

- Por lo que respecta a las características físicas de la población se busca medir su grado de funcionamiento sensorial, concretamente la vista y el oído, así como otros rasgos o características físicas como el peso, la talla y la relación entre ambos. Finalmente, se investigan las condiciones bucodentales.

1.6.4.2. Características y condiciones de uso de los servicios de salud

- Consulta médica, con detalle del lugar, motivo, especialidad, distancia o trayecto, tiempo de espera y tipo de asistencia recibida.

- Hospitalizaciones detallando la frecuencia, duración, tipo de asistencia o tratamiento, forma de ingreso, entidad pagadora, y, en su caso, lista de espera.

- Utilización del servicio de urgencias con detalle del tipo y frecuencia.

- Vacunación contra la gripe con especificación del agente motivador.

- Hábitos de consulta y atención ginecológica en las mujeres mayores de 16 años.

1.6.4.3. Hábitos de vida

- Consumo de bebidas alcohólicas con detalle de la frecuencia, cantidad consumida por tipo de bebida (personas de 16 y más años).
- Consumo de tabaco detallando la frecuencia, modalidad de consumo, cantidad e historial del hábito de consumo de tabaco (personas de 16 y más años).

- Horas diarias de sueño.
- Actividad física que requiere la actividad principal y la desarrollada en tiempo libre.
- Hábitos de nutrición.

1.6.4.4. Características específicas de la población infantil

- Tipo de lactancia en las primeras semanas y meses de vida.
- Conocimiento de determinadas vacunas.
- Tiempo dedicado a ver la televisión.

1.6.5. Fases de la encuesta

Conocidos ya los componentes del concepto de salud que se van a investigar, que es el punto de partida de la investigación, y establecido el criterio de cuáles son las variables que interesa estudiar, se ha planteado la realización del estudio en dos fases, que se identifican la primera fase con el Cuestionario de Hogar y la segunda con el Cuestionario de Adultos (personas de 16 y más años) y el Cuestionario de Menores (personas de 0 a 15 años).

En la primera fase se intenta captar a todas las personas residentes en el hogar, solicitándoles a todos sus miembros información sobre ciertas variables sociodemográficas. También se pregunta sobre quién se ocupa del cuidado de los niños y del trabajo de la casa.

En la segunda fase se recoge información de una única persona de 16 y más años seleccionada aleatoriamente dentro del hogar a través del Cuestionario de Adultos y de un único menor (personas de 0 a 15 años) de cada hogar si los hubiere, seleccionado aleatoriamente dentro del hogar a través del Cuestionario de Menores. A estas personas se les pasa una batería de preguntas referentes a los temas objeto de estudio: medida de la utilización de los servicios sanitarios y sociales, autovaloración del estado de salud, características antropométricas, limitaciones temporales de las actividades cotidianas, hábitos de vida, accidentalidad, prevalencia de enfermedades crónicas, prácticas preventivas y hábitos de nutrición.

En cuanto a la muestra, para evitar la acumulación en las provincias más pobladas, la distribución por comunidades autónomas tiene en cuenta su propio tamaño, pero se compone de una parte mínima uniforme y fija (de 120 encuestas aproximadamente según el año y tamaño total de la muestra); y otra que se asigna de forma proporcional al peso poblacional de cada provincia, en base al censo de población y vivienda. Así, el tamaño muestral ha ido de 6.400 (en 1995 y 1997) a más de 21.000 encuestas (en 1993, 2001 y 2003). Para estimar todas las características de la misma se han utilizado técnicas de reponderación, teniéndose en cuenta la edad y el sexo de cada comunidad autónoma, en base a la Unidad de Proyección Demográfica ajustada. Así pues, el diseño muestral se calificaría de, polietápico (salvo en 1993), estratificado, desagregado por comunidades autónomas, seleccionando los municipios y los individuos de forma aleatoria, y respetándose cuotas de edad y sexo de la población. La recogida de información (trabajo de campo) se ha hecho mediante entrevista personal en el domicilio del sujeto, por personal entrenado y evitándose el lenguaje técnico para facilitar al máximo su comprensión (por ejemplo, se hablaba de “receta” en lugar de “prescripción”). El cuestionario modelo, en cada uno de los años, dispone tanto de preguntas dicotómicas, como múltiples y categorizadas, pero siempre cerradas. La complementación telefónica ha sido excepcional, pero ha permitido completar y corregir datos imprecisos o incongruentes, tras un análisis preliminar.

1.6.6. Ámbito de la encuesta

1.6.5.1. Ámbito poblacional

La investigación se dirige al conjunto de personas que reside en viviendas familiares principales. Cuando una misma vivienda esté constituida por dos o más hogares, el estudio se extiende a todos ellos, pero de manera independiente para cada hogar.

1.6.5.2. Ámbito geográfico

La encuesta se realiza en todo el territorio nacional. El ámbito geográfico de las encuestas ocupa toda España, a nivel nacional y de cada una de las 17 comunidades autónomas, así como Ceuta y Melilla.

1.6.5.3. Ámbito temporal

El período de recogida de la información es trimestral.

1.6.7. Diseño de la muestra

1.6.7.1. Tipo de muestreo

El tipo de muestreo utilizado es un muestreo polietápico estratificado. Las unidades de primera etapa son las secciones censales. Las unidades de segunda etapa son las viviendas familiares principales. Dentro de ellas no se realiza submuestreo alguno, investigándose a todos los hogares que tienen su residencia habitual en las mismas. Dentro de cada hogar se selecciona a un adulto (16 o más años) para cumplimentar el Cuestionario de Adultos y en caso que hubiera menores (de 0 a 15 años) se selecciona asimismo un menor para cumplimentar el Cuestionario de Menores.

El marco utilizado para la selección de la muestra es un marco de áreas formado por la relación de secciones censales utilizadas en el Padrón Municipal de habitantes. Para las unidades de segunda etapa se ha utilizado la relación de viviendas familiares principales en cada una de las secciones seleccionadas para la muestra. Las unidades de primera etapa se agrupan en estratos de acuerdo con el tamaño del municipio al que pertenece la sección. Se consideran los siguientes estratos:

- Estrato 0: Municipios de más de 500.000 habitantes.
- Estrato 1: Municipio capital de provincia (excepto los anteriores).
- Estrato 2: Municipios con más de 100.000 habitantes (excepto los anteriores).
- Estrato 3: Municipios de 50.000 a 100.000 habitantes (excepto los anteriores).
- Estrato 4: Municipios de 20.000 a 50.000 habitantes (excepto los anteriores).
- Estrato 5: Municipios de 10.000 a 20.000 habitantes.
- Estrato 6: Municipios con menos de 10.000 habitantes.

Para cada comunidad autónoma se diseña una muestra independiente que la representa, por ser uno de los objetivos de la encuesta facilitar datos con este nivel de desagregación.

1.6.7.2. Tamaño de la muestra

Para cubrir los objetivos de la encuesta de poder facilitar estimaciones con un determinado grado de fiabilidad a nivel nacional y de comunidad autónoma, se ha seleccionado una muestra de viviendas distribuidas en secciones censales. Las secciones se seleccionan dentro de cada estrato con probabilidad proporcional a su tamaño. Las viviendas, en cada sección, con igual probabilidad mediante muestreo sistemático con arranque aleatorio. Este procedimiento conduce a muestras autoponderadas en cada estrato. Para la selección de la persona que debe cumplimentar el Cuestionario de Adultos y del menor de 16 años (si hubiera menores en el hogar) que debe cumplimentar el Cuestionario de Menores se utiliza un procedimiento aleatorio que asigna igual probabilidad a todos los adultos (en el primer caso) y a todos los menores (en el segundo caso). La muestra se ha distribuido uniformemente entre los cuatro trimestres que componen el ámbito temporal de la encuesta. Dentro de cada trimestre se ha procurado asimismo que la distribución de la muestra por semana de referencia fuera lo más homogénea posible.

1.6.7.3. Metodología estadística

El estudio se desarrolla en 2 fases. En la primera fase se utiliza el cuestionario de hogar, donde se solicita información sobre las variables sociodemográficas y también sobre datos tales como quien se ocupa de los niños, si los hay, o del trabajo de la casa. En la segunda fase se pasa el cuestionario de adultos (16 y más años), que contiene los datos fundamentales para el presente trabajo, y de menores (0 a 15 años) si los hubiera. En esta fase, de cada hogar se recoge la información de un solo adulto seleccionado aleatoriamente, y de un solo menor si lo hay.

-Ámbito poblacional y geográfico: el ámbito poblacional queda definido como el conjunto de personas que residen en viviendas familiares principales, y el ámbito geográfico es todo el territorio nacional.

-Selección de la muestra dentro de cada estrato, las secciones se seleccionan con probabilidad proporcional a su tamaño, y dentro de cada sección las viviendas se seleccionan con igual probabilidad mediante muestreo sistemático con arranque

aleatorio. Posteriormente, para la selección de las personas dentro de cada vivienda se utiliza proceso aleatorio que concede la misma probabilidad a todas. Errores de muestreo para la estimación de los errores se ha utilizado el método Jackknife.

-Método de recogida de la información. El método es la entrevista personal, que puede ser complementada con entrevista telefónica si es necesario.

En conclusion, las ENS son un medio óptimo para obtener y analizar datos sobre los hábitos de vida de la población española y relacionar la morbilidad diagnosticada de ciertas patologías con hábitos de vida. Dado que desde el año 2001 las ENS contienen bloques temáticos referentes a la frecuencia de consumo de los principales grupos de alimentos, estas encuestas constituyen una base idónea para los objetivos del trabajo que aquí se describen.

1.7. Justificación

Las enfermedades alérgicas se encuentran entre las causas más frecuentes de patología crónica, se estima que su prevalencia está aumentando desde hace años y suponen una carga adicional y cada vez en mayor grado a los costes sanitarios.

En general, se admite que la alergia se comporta como una patología que se extiende cada vez más y afecta a una proporción importante de la población.

También se da por sentado que este crecimiento aparece fundamentalmente en países industrializados con modos de vida occidentales y que la contaminación tendría un papel primordial. Además numerosos trabajos de revisión insisten en los efectos del cambio climático y del calentamiento global como factores fundamentales en la producción de alergia. Alterarían factores físicos como presión atmosférica, pluviosidad, temperatura ambiental, etc y todo ello incidiría sobre los órganos diana afectados por la patología de alergia y asma. Pero también modificarían factores biológicos, como podría ser la producción de polen por parte de las plantas y esto sería de gran trascendencia en la patogenia de las polinosis, es decir todas las enfermedades causadas por la sensibilización al polen tanto a nivel ocular y de las vías aéreas altas como de alteraciones broncoespásticas.

Sin embargo, si buscamos confirmación cuantificada de todas estas afirmaciones, nos encontramos que los datos epidemiológicos de alergia y asma en nuestro país son muy escasos en la población general española. Normalmente tenemos muestreos no representativos, de grupos de pacientes que acudían al alergólogo con la sospecha de padecer una enfermedad de origen alérgico, sirviendo para constatar las diferentes proporciones de alérgenos que causan alergia y cuales son los métodos diagnósticos y tratamientos que emplean con más frecuencia los especialistas. Además, no existen estudios que tengan en cuenta los datos agregados que se obtienen de estudios ecológicos y que pone el punto de mira en las poblaciones y no en los individuos. Por tanto, se deberían realizar estudios en nuestro país para establecer la prevalencia de enfermedades alérgicas en una muestra de población general. También es importante verificar, en nuestra zona, estas mismas necesidades y si tenemos datos de ello.

II. HIPÓTESIS Y OBJETIVOS

2.1 Hipótesis

Las enfermedades alérgicas se encuentran entre las causas más frecuentes de patología crónica, su prevalencia está aumentando desde hace años y suponen una carga adicional y cada vez en mayor grado a los costes sanitarios.

Los estudios de prevalencia de alergia son muy escasos en la población general española. Por tanto, se deberían realizar estudios en nuestro país para establecer la prevalencia de enfermedades alérgicas en una muestra de población general, así como establecer las causas y las manifestaciones de esta alergia. También es importante verificar, en nuestra zona esta misma proposición.

2.2 Objetivos

2.2.1. Objetivo general

Estimar la prevalencia de la alergia y asma a nivel nacional y por comunidades autónomas y evaluar las diferentes asociaciones con factores sociales, económicos u otros factores de salud en un periodo de 25 años. Para ello valorar la importancia de la gran cantidad de datos agregados que existen en general sobre cuestiones médicas y resaltar el escaso aprovechamiento que se hace de ellos por los profesionales.

2.2.2. Objetivos específicos

1. Determinar y extraer conclusiones de las enfermedades de alergia y asma desde los datos agregados, acumulados en las estadísticas oficiales, de las Encuestas Nacionales de Salud del Ministerio de Sanidad y del Instituto Nacional de Estadística. Recopilar los datos epidemiológicos existentes en otras publicaciones previas.

2. Determinar la prevalencia de alergia (y del asma, la cual se considera de forma independiente en las estadísticas oficiales) a nivel nacional. Además, considerar varios factores en otros apartados.

- 2.1. Conocer si realmente las series temporales nos muestran la idea preconcebida de que alergia y asma están aumentando en los últimos años.
 - 2.2. Conocer a que grupos de edad y sexo afecta con más frecuencia.
 - 2.3. Investigar si determinados factores sociales (tipo de población, condición social, económica, laboral y nivel de estudios) influyen en la prevalencia.
3. Saber si las zonas más contaminadas, como son las grandes ciudades muestran un mayor nivel de alergia y asma. Para ello, estudiar diferencias de prevalencia según tipo de población por número de habitantes.
4. Estudio ecológico por comunidades autónomas de posibles relaciones de alergia y asma con otros factores de salud como: sedentarismo, actividad física en el trabajo, hábito tabáquico, peso corporal, tipo de desayuno, etc.
5. Determinar la prevalencia de alergia y asma diseccionando también datos por comunidades autónomas, y en especial los referidos a la Comunidad Valenciana.
6. Aplicar técnicas estadísticas como son los estudios multivariantes, y sistematizar el análisis de datos para conseguir resultados valorables. El análisis multivariante en su versión estadística y su aplicación gráfica permite una visión diferente de múltiples datos clínicos para poder sacar consecuencias prácticas.

III. MATERIAL Y MÉTODO

3.1. Tipo de estudio y descripción de la estadística empleada

Se ha realizado un estudio observacional transversal ecológico utilizando fundamentalmente datos agregados anonimizados provenientes de la Encuesta Nacional de Salud, los cuales están a disposición del público tanto en la *web* del Ministerio de Sanidad como en la *web* del Instituto Nacional de Estadística.

La estadística trabaja con datos estocásticos que es lo mismo que decir aleatorios, y que bajo las mismas condiciones pueden producir resultados diferentes pero con una distribución regular si se repiten muchas veces (el ejemplo más claro serían los resultados obtenidos por sucesivos lanzamientos de un dado). Lo contrario serían resultados obtenidos con modelos deterministas, en los que una causa precede a un efecto, el cual es predecible a través de una fórmula.

La estadística es la rama de las matemáticas aplicadas que permite estudiar fenómenos cuyos resultados son en parte inciertos. Al estudiar sistemas biológicos, esta incertidumbre se debe al desconocimiento de muchos de los mecanismos fisiológicos y fisiopatológicos, a la incapacidad de medir todos los determinantes de la enfermedad y a los errores de medida que inevitablemente se producen. Así, al realizar observaciones en clínica o en salud pública, los resultados obtenidos contienen una parte sistemática o estructural, que aporta información sobre las relaciones entre las variables estudiadas. El objeto de la estadística consiste en extraer la máxima información sobre estas relaciones estructurales a partir de los datos recogidos. La estadística tiene un doble objetivo: la generación y recopilación de datos que contengan información relevante sobre un determinado problema, y en un segundo lugar, el análisis de dichos datos con que extraer información. En estadística se distinguen dos grandes grupos de técnicas: la estadística descriptiva, en la que se estudian las técnicas necesarias para la organización, presentación y resumen de los datos obtenidos, y la estadística inferencial, en la que se estudian las bases lógicas y las técnicas mediante las cuales pueden establecerse conclusiones sobre la población a estudio a partir de los resultados obtenidos en una muestra. La inferencia estadística se basa en la formación de hipótesis acerca de los parámetros y en contrastar si la hipótesis es cierta o falsa. La decisión sobre la naturaleza de la

relación entre variables es ajena a la estadística, solo es posible determinarla a partir del conocimiento del tema que se está estudiando y para la interpretación de resultados esto es fundamental (214).

3.2. Tipos de variables

El proceso seguido en el estudio estadístico de una cierta característica o variable, puede subdividirse en tres pasos sucesivos (215).

Recogidos los datos que correspondan, el primer análisis que realizaremos es el del tipo de variable que pretendemos estudiar (cualitativa/categórica o cuantitativa). Esto condicionará en gran medida su posterior tratamiento.

Tenemos variables categóricas, que son de dos tipos: las llamadas variables nominales (que son categorías sin orden) como el sexo; y las variables ordinales (que sí representan un orden, como por ejemplo el nivel de estudios). Las variables nominales pueden ser binarias o dicotómicas (por ejemplo fumador/no fumador, enfermo/sano).

En las cuantitativas, los valores de las observaciones son numéricos (cuantificables) y, en consecuencia, ordenables. A su vez las variables cuantitativas se subdividen en dos tipos; discretas si toman valores concretos y vienen dadas por números enteros; o continuas que pueden tomar todo tipo de valores de un intervalo y que se representan por números reales.

Determinado el modo de agrupamiento de las observaciones, procedemos a su recuento. Las herramientas estadísticas disponibles en general son: para dos variables categóricas las tablas de contingencia, para una variable numérica y otra categórica la comparación de medias entre grupos y para dos variables numéricas el análisis de correlación.

Si la variable es cualitativa, construyendo la tabla de frecuencias observamos los valores diferentes de la misma. Si es cuantitativa buscaremos los valores mínimo y máximo obtenidos. Si los valores de la variable numérica, al ser agrupados según las categorías de la segunda variable, constituyen clases diferentes entre sí, se puede comparar la distribución de la variable numérica entre tantas clases o grupos como tenga la variable cualitativa. Si la relación es fuerte estamos en condiciones de

predecir el valor que toma la variable numérica conociendo la categoría del individuo observado (215).

3.3. Tipos de parámetros estadísticos

En la estadística descriptiva la obtención de muy diversas conclusiones respecto de la variable estudiada, se podrá realizar con auxilio de los diferentes parámetros estadísticos (de centralización, posición, dispersión, etc) (214,215).

3.3.1. Medidas de centralización

Las medidas de centralización más importantes son:

- **Media:** es el resultado de dividir la suma de todas las observaciones entre el número de ellas. Es el centro de gravedad de la distribución.
- **Moda:** es el valor que más se repite. Será pues el valor (o valores) cuya frecuencia absoluta sea la mayor de las observadas. Si los datos se encuentran agrupados en intervalos, obtendremos el intervalo en el que se encuentra la moda (intervalo modal).
- **Mediana:** supuestas ordenadas las observaciones, es el valor de la variable que está en el centro de las mismas. Deja pues a la mitad (el 50%) de las observaciones por debajo de dicho valor. Es el punto medio en el que se muestran tanto por encima como por debajo el 50% de los datos. La media y la mediana coinciden si la distribución es simétrica; si no coinciden es preferible utilizar la mediana porque es menos sensible a datos extremos (216).

3.3.2. Medidas de posición

Las medidas de posición permiten el cálculo del valor de la variable que ocupa una cierta posición relativa respecto del conjunto total de los valores observados.

- ✓ **Cuantil o percentil:** es aquel valor que divide a la variable en dos partes. Los percentiles van de 1 a 100. Los cuantiles de 0 a 1. El percentil 12 es igual al cuantil 0,12.
- ✓ **Mediana:** percentil de orden 50.

- ✓ Cuartiles: percentiles de órdenes 25 (Cuartil 1º), 50 (Cuartil 2º) y 75 (Cuartil 3º).
- ✓ Deciles: percentiles de órdenes 10, 20, ..., 90 (Deciles 1º, 2º, ..., 9º).
- ✓ Rango, recorrido o amplitud total: Máximo -Mínimo. Con el fin de medir el mayor o menor grado de separación de las observaciones en una primera instancia se define el rango (también denominado recorrido o amplitud total), como la diferencia existente entre los valores máximo y mínimo observados (214, 215).

3.3.3. Medidas de distribución

Las medidas de distribución nos permiten conocer la homogeneidad de la distribución sin necesidad de representarlos gráficamente (216).

- Asimetría: simetría de distribución respecto a la media. Valores por debajo de 0 describen colas alargadas a la izquierda por debajo de la media, y valores superiores a 0 colas alargadas a la derecha por encima de la media. Cuando el número de valores de una distribución es mayor en la parte inferior a la media que en la superior a la misma, la distribución se muestra asimétrica a la izquierda y hacia la derecha en caso contrario.
- Curtosis: es una medida estadística que indica el grado de concentración de una distribución alrededor de su media, o grado de concentración de los valores en la región central y su influencia sobre las colas de la distribución y se divide en: mesocúrtica, leptocúrtica y platicúrtica.

3.3.4. Medidas de dispersión

Reflejan la dispersión de la distribución, como por ejemplo la separación de los datos respecto a su media (214, 215, 216).

- ❖ Desviación media: es la media de las desviaciones, o separaciones de cada una de las observaciones, respecto a la media aritmética, consideradas en valor absoluto. Sustituyendo la media por la moda o la mediana, definiremos las desviaciones medias respecto de la moda y de la mediana.
- ❖ Varianza: es la media de los cuadrados de las desviaciones o separaciones de cada una de las observaciones, respecto a la media aritmética.

- ❖ Desviación típica/desviación estándar: es la raíz cuadrada de la varianza. Con ello corregimos la necesidad, para el cálculo de la varianza, de haber elevado al cuadrado las diferencias de las observaciones con respecto a la media y así conseguir valores positivos. Esta medida de dispersión es la más característica.
- ❖ Coeficiente de variación de Pearson: es igual a la desviación típica dividida por la media. Es adimensional y adopta valores entre 0 y 1, siendo este la máxima variabilidad. Mide la representatividad de la media. Valores extremos del mismo nos llevarán a concluir que la media no es representativa, es decir, existirán valores entre las observaciones que se separan significativamente de las demás. Solo puede ser utilizado cuando los valores de la variable toman valores "normales". Es decir, no son muy elevados ni muy pequeños, ya que una media próxima a cero o muy alta darían valores nulos o infinitos al coeficiente. Si la media es representativa de las observaciones (no existen valores extremos exageradamente distanciados de la mayoría), el coeficiente de variación permite comparar la dispersión de dos series estadísticas: mayor coeficiente indica menor homogeneidad, o lo que es lo mismo, mayor dispersión o variabilidad. Compara la dispersión relativa de varias distribuciones de frecuencias.

3.3.5. Medidas de asociación

La medidas de asociación nos indican la fuerza de relación entre variables (215, 216, 217).

- ◆ Covarianza: producto de las desviaciones medias de las dos variables dividido por "n" (número de observaciones). Indica si hay asociación lineal y en que medida cambian dos variables respecto a la media.
- ◆ Coeficiente de correlación de Pearson: es la covarianza dividida por el producto de las desviaciones típicas de dos variables. Mide la intensidad de la asociación lineal. Mientras más cerca de 1 se encuentren los coeficientes de correlación más alta es la relación entre variables. Usualmente notada como R.

- ◆ Coeficiente de determinación: proporción de la varianza total que es recogida por la regresión. Refleja la bondad o cercanía del ajuste de un modelo a la variable. Toma valores entre 0 y 1, en este último caso indica mayor ajuste. Usualmente notada como R^2 .
- ◆ Matriz de correlación: es una matriz simétrica cuyos coeficientes de correlación se encuentran por encima y por debajo de la diagonal principal (formada por los resultados que toman el valor uno). Contiene los coeficientes de correlación de Pearson.

3.3.6. Análisis multivariante

En la mayoría de las investigaciones científicas es necesario analizar relaciones simultáneas entre tres o más variables. La investigación científica es un proceso iterativo, primero es necesaria la formulación explícita de las hipótesis que después han de contrastarse mediante la recogida y el análisis de los datos. Probablemente, los resultados de estos análisis sugieran una modificación de las hipótesis, siendo normal que durante este proceso se añadan y eliminen continuamente variables. La complejidad de los fenómenos analizados hace que sean muchas las variables implicadas, de allí que las investigaciones sean necesariamente multivariantes (204).

Los objetivos que persigue el análisis multivariante, podrían sintetizarse del siguiente modo: proporcionar métodos cuya finalidad es el estudio conjunto de datos multivariantes, lo que a través del análisis estadístico uni y bivalente no es posible lograr; y ayudar al investigador a tomar decisiones óptimas en el contexto en el que se encuentre teniendo en cuenta la información que aporta el conjunto de datos analizado.

Técnicas de dependencia

- Regresión lineal múltiple. Es la técnica adecuada si en el análisis hay una variable dependiente cuyo valor depende de varias variables independientes métricas o no métricas. En diversas investigaciones esta técnica tiene un papel fundamental puesto que su aplicación permite, entre otras utilidades, observar de qué manera las variables independientes logran predecir la

variable dependiente o criterio. En otras palabras, a partir del análisis de regresión será posible inferir acerca de la existencia o no de relaciones significativas entre las variables independientes y la variable criterio o respuesta, siempre dentro del marco científico establecido para la investigación en curso (205).

- **Análisis discriminante.** Esta técnica se utiliza para clasificar a distintos individuos en grupos (o poblaciones) alternativos a partir de los valores de un conjunto de variables sobre los individuos a los que se pretende clasificar. Cada individuo puede pertenecer a un solo grupo. La pertenencia a uno u otro grupo se introduce en el análisis mediante una variable categórica que toma tantos valores como grupos existan. En el análisis discriminante esta variable desempeña el papel de variable dependiente.
- **Correlación canónica.** Este análisis pretende determinar la existencia de asociación lineal entre un conjunto de variables independientes y otro conjunto de variables dependientes, de acuerdo con el esquema simplificado.
- **Análisis multivariante de la varianza (MANOVA).** En este análisis se consideran simultáneamente varias variables dependientes que supuestamente están relacionadas entre sí. Se utiliza en situaciones en las que la muestra total está dividida en varios grupos basados en una o varias variables independientes (factores) no métricas, mientras que las dependientes son métricas. Su objetivo (similar al que persigue la correlación canónica) es averiguar si hay diferencias significativas entre dichos grupos en cuanto a las variables dependientes se refiere (205).

Técnicas descriptivas o de interdependencia

En caso de encontrarse ante una situación en la que sea imposible distinguir conceptualmente entre variables dependientes e independientes, y lo que interesa es determinar cómo y por qué las variables están correlacionadas entre sí; entonces, se debe recurrir a los métodos estadísticos de interdependencia, pues son los que abordan estas cuestiones (204, 206, 207).

- **Análisis de componentes principales.** Es una técnica de reducción de datos cuyo objetivo fundamental es construir combinaciones lineales de las variables originales que expliquen la mayor parte posible de la información recogida en estas variables. Cada combinación lineal se extrae de tal forma que están incorrelacionadas con las anteriores.
- **Análisis factorial.** También es una técnica de reducción de datos, como la anterior, pero a diferencia de ésta, el interés se centra en establecer qué variables latentes (factores) pueden estar causando la correlación entre las variables observadas. Puede verse como una técnica que pretende identificar grupos de variables de tal forma que las correlaciones entre las variables de cada grupo sean superiores a las correlaciones de las variables entre los grupos. En resumen, puede decirse que ambas técnicas se usan para analizar interrelaciones entre un número elevado de variables métricas explicando dichas interrelaciones en términos de un número menor de variables denominadas factores (si son inobservables) o componentes principales (si son observables).
- **Análisis de correspondencias.** Esta técnica estadística pretende representar en un espacio multidimensional reducido la relación existente entre las categorías de dos variables no métricas. En el análisis de correspondencias, el mapa perceptual que se obtiene muestra las distancias entre los distintos niveles de dos variables, por lo que suele decirse que este análisis sirve para visualizar tablas de contingencia. Por ejemplo, si se desea saber si existe relación entre dos variables, cada una de las cuales tiene varios niveles. La tabla que se obtendría impediría distinguir con facilidad lo que se desea conocer; sin embargo, el análisis de correspondencias permitiría mostrar un mapa, probablemente bidimensional, en el cual las variables estarían tanto más cerca cuanto mayor sea la frecuencia de esa celda en la tabla de contingencia.
- **Análisis de conglomerados,** al que también se denomina comúnmente análisis *cluster*, es una técnica diseñada para clasificar distintas

observaciones en grupos (diferente del análisis factorial, debido a que la finalidad de éste radica en agrupar variables), de tal forma que:

1. Cada grupo (conglomerado o *cluster*) sea homogéneo respecto a las variables utilizadas para caracterizarlos; es decir, que cada observación contenida en él sea parecida a todas las que estén incluídas en ese grupo.
 2. Que los grupos sean lo más distintos posibles unos de otros respecto a las variables consideradas.
- Escalamiento multidimensional.

3.4. Fuentes de datos: La Encuesta Nacional de Salud

Se procedió a conseguir los datos de las Encuestas Nacionales de Salud y del INE. Son datos a disposición del público y están en formato Excel, por lo que son manejables posteriormente. El tipo de estudio es transversal de base poblacional. En cuanto a la población de estudio y fuentes de información: los datos utilizados son de los años 1993, 1995, 1997, 2001, 2003, 2006, 2009, 2011 y 2014 y 2017 y proceden del INE (Instituto Nacional de Estadística) y de la Encuesta Nacional de Salud. Se trata de encuestas de base poblacional realizadas a través de entrevistas personales en el hogar mediante un cuestionario estructurado. La población de referencia es la de 16 años o más, no institucionalizada, residente en España.

Se trataban de datos acerca de patología de asma y alergia y su afectación según determinadas condiciones:

- sexo, grupo de edad y tipo de población
- sexo, grupo de edad y nivel de estudios
- sexo, grupo de edad y nivel económico/clase social
- sexo, grupo de edad y tipo de ocupación
- sexo, grupo de edad y tipo de jornada laboral

Las Encuesta Nacional de Salud (ENS) recogen información sobre diferentes materias relacionadas con la salud, como son la morbilidad percibida, la utilización de servicios sanitarios y actividades preventivas, hábitos de vida y características personales y sociodemográficas. La forma y la cantidad de información recogida sobre estas materias se han ido modificando con los años, a medida que la

investigación epidemiológica ha ido resaltando los datos demográficos, sociológicos, de hábitos de vida, etc, de mayor importancia para el conocimiento más amplio de la salud y de la enfermedad. Los cuestionarios son los documentos fundamentales en la recogida de la información de la encuesta, y constituyen los instrumentos básicos que utiliza el entrevistador en la realización de su trabajo. Los cuestionarios de la Encuesta Nacional de Salud, incluyen un conjunto de preguntas que deben ser formuladas a todos los entrevistados de la misma manera, facilitando la toma de datos que precisa la investigación y garantizando que las respuestas a las preguntas pueden ser comparables.

El cuestionario de Hogar es un documento diseñado con la finalidad de captar a las personas del hogar para recoger información sobre las características sociodemográficas de todos sus miembros y de determinados aspectos relativos al hogar encuestado (situación económica, cuidado de los niños y trabajo de la casa), siendo por tanto el hogar la unidad de observación. El método de recogida de información para este cuestionario será siempre el de entrevista personal, complementada cuando sea necesario mediante entrevista telefónica. Como norma general, el informante del Cuestionario de Hogar debe ser el sustentador principal, salvo cuando concurra alguna circunstancia que impida la realización de la entrevista con dicha persona por razones de ausencia temporal, incapacidad o enfermedad, desconocimiento del idioma, entre otras. En estos casos podría facilitar la información otra persona del hogar, debiendo seleccionarse dicho informante por el siguiente orden de preferencia: cónyuge o pareja, padre o madre, hermano o hijo (teniendo siempre en cuenta la edad del sujeto) de la persona de referencia, otro familiar u otro miembro del hogar no vinculado a la persona de referencia por lazos de parentesco, eligiéndose en este último caso el miembro del hogar que haya residido más tiempo en el mismo. En este cuestionario se solicitan informaciones sobre el tipo de hogar y composición del mismo, así como el nivel mensual de ingresos regulares del hogar, prestaciones sociales extraordinarias recibidas por sus miembros, datos sobre el cuidado de los niños y del trabajo de la casa. Así mismo, se recoge información de todos los residentes en el hogar (miembros del hogar y otras personas residentes en el hogar) sobre edad, sexo, nacionalidad, estado civil, relación de parentesco con la persona principal y otros temas fundamentales como

nivel de estudios terminados; relación con la actividad económica, situación profesional, ocupación actividad del establecimiento; cobertura sanitaria; limitaciones para las actividades de la vida cotidiana; ayudas para realizarlas. Por ultimo, permite determinar qué personas de 16 y más años residen en el hogar, para seleccionar el sujeto de entrevista del Cuestionario de Adultos y qué personas de 0 a 15 años residen, para seleccionar el sujeto de entrevista del Cuestionario de Menores.

El Cuestionario de Adultos es un documento diseñado con la finalidad de recoger información relativa a personas de 16 y más años sobre cuatro objetivos básicos:

- Estudiar los problemas de salud y accidentalidad que pueden afectar a la población de manera importante.
- Determinar la utilización de los servicios sanitarios.
- Conocer la implantación entre la población de prácticas preventivas.
- Identificar los hábitos de vida que inciden en el estado de salud general de la población, así como las características personales y, en el caso de que el sujeto de entrevista tenga 65 años o más, proporcionar información sobre su capacidad para desarrollar determinadas actividades de la vida diaria.

Clase social ocupacional: se obtiene la clase social del individuo que trabaja o ha trabajado y del sustentador principal, y se escoge la más privilegiada de estas dos, agrupando en las clases sociales I y II (más privilegiadas), III, IV, V, VI (menos privilegiadas) de acuerdo con la adaptación española de la clasificación del *British Registrar General* (218, 219, 220, 221). En 2017 la clase social también se ha obtenido según la clasificación propuesta en 2012 por el grupo de trabajo de determinantes de la Sociedad Española de Epidemiología (SEE), adaptada para la ENS. Es una agrupación de ocupaciones, actuales o anteriores, codificadas según la Clasificación Nacional de Ocupaciones (CNO-11).

- I. Directores/as y gerentes de establecimientos de 10 o más asalariados/as y profesionales tradicionalmente asociados/as a licenciaturas universitarias.
- II. Directores/as y gerentes de establecimientos de menos de 10 asalariados/as, profesionales tradicionalmente asociados/as a

- diplomaturas universitarias y otros/as profesionales de apoyo técnico.
Deportistas y artistas.
- III. Ocupaciones intermedias y trabajadores/as por cuenta propia.
 - IV. Supervisores/as y trabajadores/as en ocupaciones técnicas cualificadas.
 - V. Trabajadores/as cualificados/as del sector primario y otros/as trabajadores/as semi-cualificados/as.
 - VI. Trabajadores no cualificados.

Nivel de estudios: sin estudios, primarios, secundarios (divididos en dos grupos) y universitarios. En 2017 el nivel educativo y de formación máximo alcanzado por la persona entrevistada se codificó según la Clasificación Nacional de Educación CNED-2014-A.

-Básico e inferior. Hasta la educación secundaria de 1ª etapa inclusive.

-Intermedio. Hasta bachillerato o enseñanzas profesionales de grado medio inclusive.

-Superior. Enseñanzas profesionales de grado superior o universitarios.

País de origen: según el país de nacimiento se agrupan los individuos en: 1) España, Unión Europea y demás países desarrollados, y 2) países considerados en vías de desarrollo (países con el índice de desarrollo humano menor que 0.9, según terminología de la Organización de Naciones Unidas).

Zonas rurales y urbanas: los individuos fueron divididos entre vivir en municipios con menos de 10.000 habitantes «rurales» o con más de 10.000 habitantes «urbanos», según criterio del Instituto Nacional de Estadística.

Apoyo social: analizado a partir del cuestionario de apoyo social de Duke que recoge dos dimensiones: el apoyo confidencial y el apoyo afectivo. La suma de los puntos de la escala se ha agrupado en cuartiles.

Situación laboral: trabajando, en baja laboral, desempleado/a, jubilado/a o pensionista, estudiando, dedicado a las actividades del hogar. Entre las personas que estaban trabajando fueron analizadas dos variables más: tipo de contrato laboral: funcionario/a o contrato de trabajo con duración indefinida; trabajo por cuenta propia; contrato temporal o sin contrato.

3.5. Grupos de datos incluidos en la ENS empleados en el presente estudio

- Tipo de estudio: Estudio transversal de base poblacional.
- Población de estudio y fuentes de información: los datos utilizados son de los años 1993, 1995, 1997, 2001, 2003, 2006, 2011 y 2014 y 2017 y proceden del INE (Instituto Nacional de Estadística) y de la Encuesta Nacional de Salud. Se trata de encuestas de base poblacional realizadas a través de entrevistas personales en el hogar mediante un cuestionario estructurado. La población de referencia es la de 16 años o más, no institucionalizada, residente en España. Para obtener una muestra representativa, se utilizó un muestreo polietápico. Las unidades de primera etapa fueron las secciones censales. Las unidades de segunda etapa, las viviendas familiares principales, y a su vez se investigaron los hogares que tenían su residencia habitual en las mismas.

Los datos escogidos llegan hasta 2017. Se decidió excluir los datos de la última Encuesta Nacional de Salud, realizada en el año 2020, en España por mostrar los resultados divididos en dos periodos por la pandemia.. La Encuesta Europea de Salud en España (EESE) 2020, realizada por el Instituto Nacional de Estadística (INE), es la parte española de la *European Health Interview Survey* (EHIS), coordinada por Eurostat y regulada por Reglamento (EC) 1338/2008 y Reglamento de la Comisión 141/2013. En este documento se muestran los resultados calculados en dos periodos, antes de la declaración del primer estado de alarma y posteriormente a este, de indicadores que han sido seleccionados atendiendo al periodo de referencia de la pregunta y al hecho de reflejar comportamientos susceptibles de modificarse debido a la situación vivida por los entrevistados. Los indicadores que se presentan en este documento no son comparables con la serie histórica debido a la estacionalidad de sus resultados. Asimismo, este hecho debe ser tenido en cuenta en la comparación de resultados entre ambos periodos. Además, no se han utilizado los datos sobre alergia y asma en infancia puesto que no van acompañados de resultados con respecto a factores sociales como condición socioeconómica, estudios, actividad laboral, etc. Incluso los diagnósticos en edad pediátrica acerca de patología respiratoria puede mostrar variación a lo largo del tiempo.

Los datos correspondían a problemas o enfermedades crónicas de larga evolución, (considerándose en este estudio las siguientes patologías: alergia, asma), padecidas en los últimos doce meses diagnosticadas por un médico y segmentadas por:

- sexo y grupo de edad
- sexo, grupo de edad y tipo de población
- sexo, grupo de edad y nivel de estudios
- sexo, grupo de edad y nivel económico/clase social
- sexo, grupo de edad y tipo de ocupación
- sexo, grupo de edad y tipo de jornada laboral

También datos relativos a determinantes de salud, por sexo, grupo de edad y comunidad autónoma: índice de masa corporal, consumo de tabaco, consumo de alcohol, ejercicio físico en el trabajo, sedentarismo.

Las tablas fueron convenientemente depuradas para poder realizar el análisis estadístico.

3.6. Estadística descriptiva realizada en este estudio

-Análisis estadístico:

En primer lugar, se realizó un estudio descriptivo de las muestras elegibles. Todos los análisis estadísticos se realizaron utilizando el paquete estadístico R versión 1.4.2. y el Excel version 2016 Microsoft Excel versión 16.0 (Microsoft Corporation. *From: Office.microsoft.com/excel*).

R es un sistema para análisis estadísticos y gráficos creado por Ross Ihaka y Robert Gentleman (Ihaka R. & Gentleman R. 1996. *R: a language for data analysis and graphics. Journal of Computational and Graphical Statistics* 5: 299–314). R tiene una naturaleza doble de programa y lenguaje de programación y es considerado como un dialecto del lenguaje S creado por los Laboratorios AT&T Bell. R se distribuye gratuitamente bajo los terminos de la GNU (*General Public Licence*); su desarrollo y distribución son llevados a cabo por varios estadísticos conocidos como el equipo central de la *R Foundation for Statistical Computing* (cran.r-project.org). (información: <http://www.gnu.org/>; <http://cran.r-project.org>). R es un lenguaje de programación muy flexible orientado a la estadística computacional, el análisis de

datos y el desarrollo de gráficos. Periódicamente se añaden paquetes con nuevas funciones que se van incorporando a través de la red CRAN (*Comprehensive R Archive Network*) que es un repositorio web utilizado por los usuarios para la distribución de trabajos de forma gratuita (222).

-Determinación de las variables a analizar estadísticamente, tanto sanitarias (por ejemplo datos de prevalencia, que es la proporción de casos de una población que presenta un determinado rasgo o enfermedad) como de otro tipo (socio-demográficas, económicas, etc). En todos los casos se efectuaron las medidas de centralización, dispersión, distribución y asociación expresadas en diferentes tablas y en diagramas (223).

- Medidas de centralización (224). Las medidas de centralización utilizadas han sido: media, mediana.
- Medidas de distribución. Se han utilizado: asimetría, curtosis.
- Medidas de dispersión: desviación media, varianza, desviación típica/desviación estándar, coeficiente de variación de Pearson.
- Medidas de asociación: covarianza, coeficiente de correlación de Pearson, coeficiente de determinación, matriz de correlación.

-Gráficas de todos los factores analizados.

-Estudio de normalidad y significatividad (223)

- Para contrastar la normalidad, usamos el test de Shapiro-Wilk.
- Para contrastar la homocedasticidad (o igualdad de varianzas), usamos el test F de Fisher, el test de Barlett y el test de Levene.
- Para verificar la significatividad se han empleado el test de Student y el test de Welch.

-Utilizando como fuentes de datos las referidas en apartados anteriores y de los años 1993, 1995, 1997, 2001, 2003, 2006, 2011 y 2014 y 2017, se practicó una depuración de los mismos para poder ser empleados estadísticamente. Se realizó:

1. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años. Según grupos de edad en años (15-24, 25-44, 35-44, 45-64, 65 y más, en algunos casos más de 75 años) y sexo (cifras totales en ambos sexos, varones y mujeres).

2. Estadística descriptiva de la evolución de la prevalencia de alergia con respecto al tipo de población. Según tipo de población (menos de 10.000 habitantes, entre 10.000 y 50.000, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes) junto con grupos de edad y sexo anteriores.

3. Estadística descriptiva de la prevalencia de alergia según nivel de estudios (sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado).

4. Estadística descriptiva de la prevalencia de alergia según nivel económico (clases sociales I, II, III, IV, V y VI; siendo la I la más alta y la VI la más baja).

5. Estadística descriptiva de la prevalencia de alergia según tipo de ocupación (ocupado, parado, inactivo).

6. Estadística descriptiva de la prevalencia de alergia según tipo de jornada laboral (partida, continua, reducida y rotatoria).

También se efectuó:

1. Estadística descriptiva de la evolución de la prevalencia de asma en 10 años. Según grupos de edad en años (15-24, 25-44, 35-44, 45-64, 65 y más, en algunos casos más de 75 años) y sexo (cifras totales en ambos sexos, varones y mujeres).

2. Estadística descriptiva de la evolución de la prevalencia de asma con respecto al tipo de población. Según tipo de población (menos de 10.000 habitantes, entre 10.000 y 50.000, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes) junto con grupos de edad y sexo anteriores.

3. Estadística descriptiva de la prevalencia de asma según nivel de estudios (sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado).

4. Estadística descriptiva de la prevalencia de asma según nivel económico (clases sociales I, II, III, IV, V y VI; siendo la I la más alta y la VI la más baja).

5. Estadística descriptiva de la prevalencia de asma según tipo de ocupación (ocupado, parado, inactivo).

6. Estadística descriptiva de la prevalencia de asma según tipo de jornada laboral (partida, continua, reducida y rotatoria).

En todos los casos, tanto para la alergia como para el asma, se realizó:

- Estadística descriptiva con medidas de posición central (media, mediana), medidas de dispersión (desviación típica, varianza, coeficiente de variación), medidas de distribución (rangos, valores máximo y mínimo, coeficiente de asimetría y curtosis).
- Aplicación de test de significatividad a los grupos inicial y final.
- La estadística descriptiva se ha traducido en tablas:
 - Tablas con porcentaje de prevalencia comparando grupos de edad, sexo, tipo de población, nivel de estudios, nivel económico, ocupación y tipo de jornada laboral.
 - Tablas con resumen de parámetros de estadística descriptiva.
- La estadística descriptiva se ha traducido en gráficos:
 - Diagrama de líneas con series temporales según los grupos mencionados.
 - Diagramas de porcentaje de prevalencia comparando grupos de edad, sexo, tipo de población, nivel de estudios, nivel económico, ocupación y tipo de jornada laboral.
 - Representación multivariante Biplot.

3.7. Gráficos y su descripción

Podremos visualizar los resultados de forma gráfica con el diagrama estadístico apropiado (224). La norma que hemos de seguir en la construcción de un gráfico estadístico es siempre que la zona que identifica a cada valor será proporcional a su frecuencia:

- Diagramas de barras para variables cualitativas o cuantitativas no agrupadas en intervalos, sobre un eje (normalmente el horizontal) marcamos los valores de la variable, dibujando sobre cada uno de ellos una barra cuya longitud sea

proporcional a la frecuencia o proporción que se esté visualizando. Se utilizan para descripción y comparación de categorías y de distribuciones.

- Histogramas: representativo de las variables agrupadas en intervalos. Sobre el eje horizontal marcamos los distintos intervalos, dibujando sobre cada uno de ellos un rectángulo cuya área sea proporcional a la frecuencia que se esté visualizando (si todos los intervalos tienen la misma amplitud, nos bastará con que la altura de los rectángulos sea proporcional a las frecuencias).
- Polígonos de frecuencias: si la frecuencia representada no es acumulada, enlazamos los puntos medios de los extremos superiores de los rectángulos, y para frecuencias acumuladas, el polígono de frecuencias se obtiene de la forma indicada en el gráfico. Si la variable representada es cuantitativa, enlazando los extremos de las barras obtendremos el polígono de frecuencias, denominado perfil ortogonal para cualitativas ordenables.
- Diagramas de sectores: utilizable en cualquier tipo de variable. Dividimos el círculo en sectores circulares, de modo que la amplitud de cada sector, sea proporcional a la frecuencia. Junto a cada sector, se suele indicar el valor representado. Es aconsejable la expresión de las amplitudes de los sectores en porcentajes (%). Si la variable es cualitativa se puede emplear el diagrama de sectores.
- Diagramas de áreas: representativo de las variables cuantitativas, equivale a la representación independiente de los polígonos de frecuencias (descritos en los diagramas de barras e histogramas). Indica la evolución de los valores de la variable, consistiendo en la visualización del área encerrada bajo el polígono de frecuencias. Para ello, se conecta dicho polígono con el eje de la variable (el horizontal en el gráfico), tanto a la izquierda del primer valor como a la derecha del último.
- Diagramas de puntos: para cada categoría de la variable independiente situada en el eje horizontal se dibuja el valor de la variable dependiente para todos los sujetos del grupo.
- Diagramas de cajas: contiene en la caja el 50% de los valores (límites primer y tercer cuartil), con la mediana en el centro de la caja y los límites máximo y mínimo saliendo de la caja, estando los valores atípicos sin contacto con los

elementos citados. Se usa para variables cuantitativas sobre todo continuas. Resumen bien la distribución de una variable.

- Curva de densidad: la altura de la curva representa la frecuencia de observaciones para una puntuación determinada. El eje vertical representa porcentajes y el eje horizontal la escala de la variable.
- Diagramas de líneas: para análisis temporales.
- Diagrama de dispersión: para observar la distribución conjunta de dos variables, siendo ambas continuas (tanto la independiente como la dependiente).
- Representación multivariante Biplot.

Dada la complejidad del tamaño de las muestras se usaron también técnicas de gráficos para datos multivariantes (biplot) (225).

Un biplot es una representación gráfica de datos multivariantes. De la misma manera que un diagrama de dispersión muestra la distribución conjunta de dos variables, un biplot representa tres o más variables. El biplot aproxima la distribución de una muestra multivariante en un espacio de dimensión reducida, normalmente bidimensional, y superpone sobre la misma representaciones de las variables sobre las que se mide la muestra. Las representaciones de las variables son normalmente vectores, y coinciden con las direcciones en las que mejor se muestra el cambio individual de cada variable. El prefijo “bi” se refiere a la superposición, en la misma representación, de individuos y variables. Los biplots son útiles para describir gráficamente los datos o para mostrar los resultados proporcionados por modelos más formales. La forma más sencilla del biplot es un diagrama de dispersión en el que los puntos representan a los individuos, y los dos ejes a las variables. Cuando tenemos tres o más variables las cosas se complican y, en general, las distribuciones multivariantes son difíciles de visualizar. Desde el punto de vista del usuario, los biplots serán importantes porque su interpretación se basa en conceptos geométricos sencillos, que forman parte de la cultura matemática de los potenciales usuarios, a saber, la similitud entre individuos es una función inversa de la distancia entre los mismos, sobre la representación biplot. En determinados tipos, las longitudes y los ángulos de los vectores que representan a las variables, se interpretan en términos de variabilidad y covariabilidad respectivamente. Las

relaciones entre individuos y variables se interpretan en términos de producto escalar, es decir, en términos de las proyecciones de los puntos “individuo” sobre los vectores “variable” (226).

Los componentes principales son proyecciones que muestran donde los datos están más extendidos y presentan una mayor varianza. Un biplot es una generalización del diagrama de dispersión simple de dos variables y permite mostrar gráficamente información sobre muestras y variables de una matriz de datos. Si graficamos mediante un círculo de correlaciones vemos que si las variables se alejan del centro hacia el lado positivo y están próximas una a otra están positivamente correlacionadas. Si están en lados opuestos al centro están negativamente correlacionadas. Si adoptan una posición ortogonal entre ellas no están correlacionadas (227).

En cuanto a las observaciones, si están próximas se interpretan en términos de similitud. El ángulo se interpreta en términos de correlación, concretamente el coseno del ángulo que forman dos rectas que representan a las variables que se cruzan, da la correlación de resultados; si las rectas son perpendiculares el coseno es igual a 0 y las variables no están relacionadas; y cuando el ángulo se acerca a los 180° la relación es alta. Las proyecciones de las observaciones sobre las variables permiten obtener los valores de la muestra de datos.

La gráfica de influencias dibuja los coeficientes de una variable para el primer componente *versus* los coeficientes del segundo componente. Nos permite identificar que variables tienen mayor efecto en cada componente. Las influencias pueden ir de -1 a 1. Las influencias que se aproximan a -1 y 1 indican que la variable afecta considerablemente al componente. Las influencias cercanas a 0 indican que la variable tiene poca influencia en el componente. La abundancia de observaciones se obtiene por las proyecciones ortogonales sobre los vectores (228).

La correlación entre variables viene dada por el ángulo existente entre vectores, siendo la mayor correlación cuando el ángulo es más pequeño y los vectores se superponen.

En el gráfico biplot lo más importante está en las direcciones y la longitud de las variables.

3.8. Estadística para verificar la normalidad y detectar la significatividad

Una vez determinados los dos pasos previos que son la verificación o no de la distribución siguiendo la normalidad y el comprobar la homocedasticidad (la homocedasticidad es la homogeneidad de varianzas), aplicamos los test de significatividad (por ejemplo por comparación de medias) (222, 223).

Todos los análisis estadísticos descriptivos se realizaron utilizando el paquete estadístico R versión 1.4.2. y con el programa Excel.

- ◆ Los tests de normalidad son:
 - Kolmobarow si hay mas de 50 casos que nos determina el grado de aportación de los datos a una distribución teórica.
 - Shapiro si hay menos de 50 casos.
- ◆ Los test de homocedasticidad son:
 - El test F de Fisher, el test de Levene y el test de Barlett.
- ◆ Los test de correlación son:
 - Test de correlación (en distribuciones paramétricas Pearson, en no paramétricas Spearman).
- ◆ Los test de significatividad son:
 - Test no paramétricos: no se cumple alguno de los requisitos de normalidad y/o homocedasticidad.
 - Dos muestras: se aplica el test Wilcoxon-Mann-Whitney. En ocasiones se pueden ver como test de Wilcoxon (comparar medianas de 2 o más grupos apareados) o test de U-Mann-Whitney (comparar medias de dos muestras independientes), por separado.
 - Más de dos muestras: se aplica el test de Kruskal-Wallis que compara medianas de 3 o más grupos independientes.
 - Test paramétricos: las muestras son normales y homocedásticas. Dependiendo del número de muestras, se usa un test u otro:
 - Variable cualitativas: test de Chi cuadrado.
 - Variables cuantitativas con cualitativas, o cuantitativa con cuantitativa:

- T de Student (compara medias de una muestra con respecto a la media habitual de la población a la que pertenece, o bien 2 muestras de datos apareados o independientes). Se aplicaría el test Welch (T test) en el caso de que las varianzas de la muestra sean diferentes.
- ANOVA para más de 2 categorías de datos.

En estadística, *significante* no significa que sea importante ni siquiera que sea verdadero, lo que quiere decir es que el investigador no lo considera por azar y que se puede repetir. En cuanto a nivel de significación estima el porcentaje de riesgo de confundirnos que estamos dispuestos a asumir y varía entre 0 y 100.

El contraste de hipótesis es el que sigue:

- **H₀**: ambos grupos presentan igualdad y no tienen diferencias significativas. Hipótesis nula, es la hipótesis de partida o inicial que enuncia que las diferencias no son significativas.
- **H₁**: ambos grupos tienen diferencias significativas.

El error tipo I es rechazar H₀ cuando es verdadero y el error tipo II es aceptar H₀ cuando es falso. Los test de significación son pruebas que se realizan para contrastar la información empírica con la teórica. Para ello se fija un umbral de corte a partir del cual los valores obtenidos en la muestra empírica se consideren poco probables. La significación estadística se acuerda habitualmente por un valor de $p < 0,05$.

3.9. Análisis de conglomerados y de componentes principales

Dada la complejidad del tamaño de los datos se procedió al análisis multivariante. Son métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos con varias variables medidas. Permiten un mejor entendimiento del fenómeno objeto de estudio, obteniendo información que los métodos univariantes y bivariantes son incapaces de conseguir. En el caso de que se elija un método de interdependencia no distinguen entre variables dependientes e independientes y su objetivo consiste en identificar qué variables están relacionadas, cómo lo están y por qué (229).

Todos los análisis estadísticos descriptivos se realizaron utilizando el paquete estadístico R versión 1.4.2.

3.9.1. Análisis de conglomerados

Es una técnica de agrupamiento que clasifica las unidades de la matriz de datos en conglomerados (grupos internamente homogéneos y heterogéneos entre ellos) con el empleo de diferentes índices de similaridad o de distancia. El análisis *cluster* es una técnica multivariante cuya idea básica es clasificar objetos formando grupos/ conglomerados (*cluster*) que sean lo más homogéneos posible dentro de sí mismos y heterogéneos entre sí (230). Esta técnica no supervisada se basa en identificar grupos en los datos de tal manera que todos los datos del grupo (conglomerado) son datos con características similares mientras que los datos de los otros grupos son diferentes. Agrupa a un conjunto de individuos o casos en grupos, también denominados conglomerados, con el siguiente criterio: los individuos dentro del conglomerado serán lo más parecidos posible (máxima homogeneidad interna). Los individuos de grupos diferentes serán lo más diferentes posible (máxima heterogeneidad entre grupos) (231).

El análisis *cluster* es una tarea de clasificación. La clasificación es uno de los objetivos fundamentales de la ciencia y en la medida en que el análisis *cluster* nos proporciona los medios técnicos para realizarla, se nos hará imprescindible en cualquier investigación. Como técnica de agrupación de variables, el análisis *cluster* es similar al análisis factorial. Como técnica de agrupación de casos, el análisis *cluster* es similar al análisis discriminante. Pero mientras que el análisis discriminante se centra en la agrupación de variables, es decir efectúa la clasificación tomando como referencia un criterio o variable dependiente (los grupos de clasificación), el análisis *cluster* se centra en agrupar objetos, es decir permite detectar el número óptimo de grupos y su composición únicamente a partir de la similaridad existente entre los casos; además, el análisis de *cluster* no asume ninguna distribución específica para las variables.

Estas clasificaciones pueden ser jerárquicas (ascendente o descendente) en las que no se establece *a priori* el número de grupos, o no jerárquicas en las que se parte de

un número dado de grupos según hipótesis o cálculos previos de grupos de objetos similares. Se emplea el escalamiento multidimensional, técnica que permite la reducción de datos, al tratar que la distancia entre los objetos en el nuevo espacio dimensional reducido sea lo más parecida a las distancias del espacio original, tal como la construcción de un mapa a menor escala. Surge ante la necesidad de diseñar una estrategia que permita definir grupos de objetos homogéneos. Este agrupamiento se basa en la idea de distancia o similitud entre las observaciones y la obtención de dichos *clusters* depende del criterio o distancia considerados. Es decir, el número de *clusters* depende de lo que consideremos como similar.

Dentro de las jerárquicas destacaremos el algoritmo dendrograma. El dendrograma es un gráfico usado en el procedimiento jerárquico que permite visualizar el proceso de agrupamiento de los *cluster* en los distintos pasos, formando un diagrama en árbol. Es una gráfica bidimensional en que en el eje Y se muestran las variables y en el eje X la distancia a la que se encuentra una variable de otra. Da una idea visual de la proximidad entre *clusters* y ayuda a decidir cuántos grupos formar. El dendrograma es una valiosa herramienta visual que puede ayudar a decidir el número de grupos que podrían representar mejor la estructura de los datos teniendo en cuenta la forma en que se van anidando los *cluster* y la medida de similitud a la cual lo hacen. Cortando el gráfico con un segmento perpendicular a las ramas se obtiene una partición con un número de grupos igual a las ramas cortadas. El número de *cluster* o clases en que deseamos agrupar los datos nos llevará a cortar el dendrograma y ver a qué nivel de similitud se da dicho agrupamiento.

Dentro de las no jerárquicas destacaremos el algoritmo K-Means. El objetivo de K-Means es claro: agrupar observaciones similares para descubrir patrones que a simple vista se desconocen. Para conseguirlo, el algoritmo busca un número fijo (k) de conglomerados. Cada conglomerado tiene asociado un centroide. Este número k es un hiperparámetro del algoritmo. Representa el número de centroides (centro geométrico del conglomerado) que queremos encontrar.

El análisis de conglomerados es una de una de las herramientas más utilizadas para ver segmentaciones de los datos y tener una intuición de cómo están estructurados los datos. Para poder diferenciar grupos dentro de una misma población de

pacientes se realizó análisis *cluster*. Las enfermedades seleccionadas fueron alergia y asma.

Los conglomerados se efectuaron tomando los siguientes datos:

- Según grupos de edad en años (15-24, 25-44, 35-44, 45-64, 65 y más) y sexo (cifras totales, varones y mujeres)
- Según tipo de población: menos de 10.000 habitantes, entre 10.000 y 50.000, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes. Con grupos de edad y sexo.
- Según nivel de estudios, grupos de edad y sexo.

Escogimos estos apartados porque los valores eran más homogéneos y las distribuciones seguían patrones de normalidad.

3.9.2. Análisis de componentes principales

Se usa para variables cuantitativas. El ACP (análisis de componentes principales) y el análisis del factor común (AFC) forman parte de un conjunto de técnicas multivariadas denominadas de interdependencia, ya que analizan la relación mutua entre un conjunto de variables. Su finalidad principal, por lo tanto, no es el análisis de relaciones causales, sino la agrupación de variables, en función de la variabilidad que cada variable comparte con otras variables (varianza o covarianza). Se busca la síntesis de la información proporcionada por cierto grupo de variables observadas, en un número inferior de variables no observadas (factores comunes o componentes principales, depende de la variedad analítica que se realice), con la menor pérdida de información posible. Dichas variables no observadas (o latentes), darían cuenta de conceptos no observables directamente, que engloban lo que tienen en común las variables observadas (232).

Las técnicas permiten reducir la dimensionalidad de los datos, con lo que se pueden descartar atributos irrelevantes para la toma de decisiones, que conspiran además en contra de la eficiencia de los algoritmos propios de las técnicas empleadas a *posteriori*. Las nuevas características encontradas son combinación lineal de las anteriores y se les llama componentes si se forman por variables observadas y factores si se forman por variables no observadas.

Esta serie menor de variables latentes (factores o componentes) se caracteriza por aglutinar variables empíricas que están bastante correlacionadas entre sí y escasamente correlacionadas con aquellas variables empíricas que conforman otra estructura latente (o dimensión del concepto que se analice).

Esto implica que la "no correlación" entre grupos de variables es una propiedad importante. Significa que los indicadores miden "dimensiones" diferentes en los datos. La técnica nos es útil, por tanto, para una serie de objetivos:

- Reducir la información de una matriz de correlaciones a partir de la construcción de funciones lineales.
- Descifrar patrones de dependencia a partir del análisis de correlaciones múltiples.
- Identificar dimensiones que representen esquemas conceptuales de análisis.
- Validar la construcción de instrumentos de medida, particularmente escalares.

Los componentes principales son las combinaciones de las variables que explican la varianza de los datos. El análisis de componentes principales (ACP) se caracteriza por analizar la varianza total del conjunto de variables observadas. De ellas trata de determinar las dimensiones básicas (o "componentes") que las definen. En el análisis de factor común (AFC) el estudio de las interrelaciones entre las variables se restringe, en cambio, a la varianza común (o covarianza), es decir, a la búsqueda de un número reducido de "factores" que expresen lo que es "común" al conjunto de variables observadas (230).

Las técnicas antes mencionadas pueden ser de amplio uso en la práctica, ya que en la clínica es frecuente emplear, como variables, múltiples determinaciones de laboratorio (variables cuantitativas) con la cual se pueden hacer investigaciones y análisis multivariados en dependencia de lo que se desea probar. Por ejemplo, en un estudio donde se quiere evaluar la eficacia de cierto tratamiento para el asma, se pueden medir un considerable número de variables que se corresponden con estas determinaciones del laboratorio clínico y funcional respiratorio, y si se tiene interés en reducir su número, en este caso se puede emplear un análisis de componentes principales (233).

En esquema:

Análisis de componentes principales:

- Técnica de reducción de datos.
- Objetivo: construir combinaciones lineales de las variables iniciales que expliquen la mayor parte de la información contenida en esas variables. Esas combinaciones se denominan componentes principales, están incorrelacionados y cada componente sucesivo explica menos varianza.

Análisis factorial:

- Técnica de reducción de datos.
- Objetivo: establecer qué causas latentes (factores) causan la correlación entre las variables observadas.

El primer paso es comprender cómo las variables del conjunto de datos de entrada varían de la media con respecto a la otra, o en otras palabras, para ver si hay alguna relación entre ellas. Porque a veces, las variables están altamente correlacionadas de tal manera que contienen información redundante. Así, con el fin de identificar estas correlaciones, calculamos la matriz de covarianza. La matriz de covarianzas es una matriz simétrica que tiene como entradas las covarianzas asociadas con todos los pares posibles de las variables iniciales.

En realidad es el signo de la covarianza lo que importa:

Si es positivo, las dos variables aumentan o disminuyen juntas (correlacionadas).

Si es negativo, una aumenta cuando la otra disminuye (inversamente correlacionada).

Geométricamente hablando, los componentes principales representan las direcciones de los datos que explican una cantidad máxima de varianza, es decir, las líneas que capturan la mayor parte de la información de los datos. La relación entre la varianza y la información aquí, es que, cuanto mayor es la varianza llevada por una línea, mayor es la dispersión de los puntos de datos a lo largo de ella. En pocas palabras, hay que pensar en los componentes principales como nuevos ejes que proporcionan el mejor ángulo para ver y evaluar los datos, de modo que las diferencias entre las observaciones sean mejor visibles. Como hay tantos componentes principales como variables en los datos, los componentes principales

se construyen de tal manera que el primer componente principal representa la mayor varianza posible en el conjunto de datos.

Los vectores propios de la matriz de covarianza son en realidad las direcciones de los ejes donde hay más varianza (más información) y que llamamos componentes principales. Y los denominados autovalores son simplemente los coeficientes unidos a autovectores, que dan la cantidad de varianza llevada en cada componente principal.

Análisis de componentes principales:

Basándose en los datos de prevalencia obtenidos de alergia y asma anteriormente citados en la estadística descriptiva.

- Según grupos de edad en años (15-24, 25-44, 35-44, 45-64, 65 y más) y sexo (cifras totales, varones y mujeres)
- Según tipo de población (menos de 10.000 habitantes, entre 10.000 y 50.000, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes) junto con grupos de edad y sexo anteriores)
- Según nivel de estudios, grupos de edad y sexo

Escogimos estos apartados porque los valores eran más homogéneos y las distribuciones seguían patrones de normalidad.

3.10. Análisis de actividad física y relación con alergia y asma

Se trataba de relacionar la actividad física tanto en el trabajo como en el ocio así como el sedentarismo con la alergia y el asma. Para ello, tomamos variables relacionadas con estos aspectos y como observaciones tomamos todas y cada una de las comunidades autónomas para las cuales existen datos cruzados de las variables reseñadas. Los datos se tomaron de dos momentos: 2011 y 2017.

Las variables tomadas fueron: actividad física intensa, actividad física moderada actividad física ligera, sedentarismo en ocio, no sedentarismo en ocio, trabajo

realizado sentado (sedentario), trabajo realizado de pie, trabajo realizado caminando, trabajo realizado con gran esfuerzo, asma y alergia.

Se realizó: estadística descriptiva, análisis de normalidad de distribución de la muestra, matriz de correlaciones y test de correlación de Pearson, análisis de componentes principales con gráfico de sedimentación, conglomerados/*Clusters* con gráfica.

3.11. Análisis de factores de salud y relación con alergia y asma

Se trataba de observar la relación entre alergia y otros factores de salud pero que no son patologías, ya que se considera que la alergia es una reacción exagerada ante agentes externos que no son nocivos para la salud pero que los pacientes sanos los identifican como tales.

Se ha escogido variables en relación con el peso corporal, hábito tabáquico, cómo se realiza el trabajo en relación con el esfuerzo físico, el sedentarismo en el ocio y el tipo de desayuno.

Se toman como observaciones cada una de las comunidades autónomas como si representaran individuos ya que existen datos agregados al respecto para todas ellas que se pueden cruzar para todas y cada una de las variables.

Las variables tomadas fueron: peso insuficiente, normopeso, sobrepeso, obesidad fumador diario, fumador ocasional, ex-fumador, nunca ha fumado, sedentarismo en ocio, no sedentarismo en ocio, trabajo realizado sentado, trabajo realizado de pie, trabajo realizado caminando, trabajo realizado con gran esfuerzo, desayuno solo líquido, desayuno líquido y cereales/tostada, asma y alergia.

Las variables manejadas según referencia del Ministerio de Sanidad eran:

- En cuanto al peso se usó el Índice de Masa Corporal (IMC): se define como la relación entre el peso del individuo (expresado en kilogramos) y el cuadrado de la talla (expresada en metros). $IMC = \text{kg}/\text{m}^2$. A partir de este índice, se define, para la población de 18 y más años:
 - Peso insuficiente ($IMC < 18,5 \text{ kg}/\text{m}^2$).
 - Normopeso ($18,5 \text{ kg}/\text{m}^2 \leq IMC < 25 \text{ kg}/\text{m}^2$).

- Sobrepeso ($25 \text{ kg/m}^2 \leq \text{IMC} < 30 \text{ kg/m}^2$).
- Obesidad ($\text{IMC} \geq 30 \text{ kg/m}^2$).
- Sedentarismo: no hace ejercicio en el tiempo libre, ocupando el tiempo de manera casi completamente sedentaria (leer, ver la televisión, ir al cine, etc.).
- Actividad física moderada o intensa. El cuestionario recoge el tiempo destinado en los últimos 7 días a 3 tipos de actividad física (intensa, moderada o caminar), -siempre que sea al menos 10 minutos seguidos-, en el trabajo, la casa, los desplazamientos y el tiempo libre de descanso, ejercicio o deporte. El nivel de actividad física bajo constituye una estimación de la población de 15 a 69 años que no cumple las recomendaciones de la OMS (al menos 150 minutos semanales de actividad física moderada o equivalente).
- Esfuerzo físico en el trabajo: trabajo sentado, de pie, caminando, trabajo con gran esfuerzo físico.
- Tipo de desayuno

Todos los datos fueron obtenidos de las Encuestas Nacionales de Salud y se practicó una depuración de tablas para poder ser empleadas estadísticamente.

Todos los análisis estadísticos se realizaron utilizando el paquete estadístico R versión 1.4.2. y con el programa Excel.

Se realizó: resumen de estadística descriptiva, análisis de normalidad de distribución de la muestra, matriz de correlaciones y test de correlación de Pearson, análisis de componentes principales con gráfico de sedimentación, conglomerados/*clusters* con gráfica.

3.12. Comparativa entre comunidades autónomas de la prevalencia de alergia y asma en 2011 y 2017.

Se realizó un análisis gráfico de los datos mediante diagramas de barras representando el porcentaje de prevalencia en cada una de las comunidades autónomas de alergia y asma.

Se graficaron mapas para situar, según porcentaje, la situación de cada comunidad autónoma, diferenciando las que han aumentado de prevalencia tanto como las que han disminuido, así como las que mantienen cifras similares.

IV. RESULTADOS

4.1. ESTADÍSTICA DESCRIPTIVA

4.1.1. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años

Los resultados recogidos van desde 1993 a 2017 con sucesivas mediciones en los años 1993, 1995, 1997, 2001, 2003, 2006, 2011, 2014 y 2017. En 1993 se declaraban alérgicos diagnosticados por un médico un 6,25% y en 2017 un 15,28% (crecimiento de un 59,09 con respecto al máximo de 2017).

En todos los grupos que figuran en la tabla 1 el aumento ha sido constante en el tiempo y cuando llegamos a 2017 las cifras de prevalencia son, en todos los nichos por edad y sexo, claramente superiores a los de los años anteriores.

La tabla 1 nos muestra la prevalencia por sexo y grupos de edad.

Tabla 1. Clasificación de la evolución de la prevalencia de alergia en España desde 1993 a 2017, tanto por cifras totales de ambos sexos, como por separado en varones y mujeres añadiendo los grupos de edad y mostrando los datos de cada uno de los años de medición.

		1993	1995	1997	2001	2003	2006	2011	2014	2017
Ambos sexos	Total	6,25	8,05	7,95	8,03	9,83	10,54	10,74	13,37	15,28
	15-24 años	7,95	11,79	10,66	10,63	13,55	13,82	13,20	13,80	19,05
	25-34 años	6,41	8,69	7,97	7,98	11,89	12,24	12,16	16,95	17,86
	35-44 años	5,99	5,47	6,97	7,50	8,85	11,41	11,92	15,28	16,37
	45-54 años	5,46	7,18	6,90	8,05	8,70	8,85	10,18	13,55	14,64
	55-64 años	5,98	7,01	7,06	7,66	8,64	9,79	9,60	12,01	14,04
	65-74 años	5,16	7,64	7,11	7,35	8,01	7,57	9,04	10,53	12,74
	75 y más años	5,34	6,08	6,91	5,61	7,06	7,70	7,33	8,24	11,44
Varones	Total	5,32	6,83	6,18	6,63	8,72	8,91	9,66	11,25	13,23
	15-24 años	7,61	11,78	9,25	10,83	14,35	13,33	13,52	12,99	18,97
	25-34 años	5,79	8,19	7,27	7,93	11,49	11,20	11,93	14,03	14,84
	35-44 años	4,70	3,89	4,72	5,24	8,42	9,44	11,35	13,19	15,88
	45-54 años	3,64	4,93	4,44	5,51	6,32	6,19	7,94	11,09	12,62
	55-64 años	4,71	5,37	4,71	4,69	6,14	7,25	7,02	9,71	10,51
	65-74 años	3,64	5,34	4,18	4,90	4,88	5,73	7,03	7,19	9,52
	75 y más años	5,17	2,94	5,20	4,62	4,44	5,82	5,70	6,63	7,33
Mujeres	Total	7,14	9,19	9,60	9,34	10,90	12,10	11,77	15,38	17,22
	15-24 años	8,33	11,80	12,20	10,43	12,73	14,34	12,86	14,65	19,14
	25-34 años	7,07	9,20	8,66	8,03	12,30	13,36	12,39	19,88	20,86
	35-44 años	7,32	7,00	9,22	9,72	9,28	13,45	12,52	17,45	16,86
	45-54 años	7,26	9,23	9,20	10,52	11,06	11,48	12,39	16,00	16,66
	55-64 años	7,19	8,55	9,21	10,40	10,99	12,18	12,05	14,21	17,42
	65-74 años	6,40	9,11	9,13	9,13	10,48	9,00	10,78	13,47	15,65
	75 y más años	5,43	8,65	8,26	6,37	8,84	9,04	8,40	9,30	14,16

La idea de que la alergia ha aumentado de forma imparable en los últimos años siempre ha estado ahí y se admitía tácitamente, pero con la tabla 1 se ve de forma nítida que tenemos unos datos cuantitativos apoyando este aumento. En varones la cifra ha pasado de 5,32% a 13,23% y en mujeres de 7,14% a 17,22%.

El grupo de edad que está entre 15-24 años es el que destaca todos los años con cifras por encima del resto de grupos de edad. Vemos que en su cifra total (es decir, teniendo en cuenta ambos sexos) ha pasado de 7,95% a 19,05%. También, es destacable el grupo de 25-34 años en mujeres que alcanza el pico más alto de toda la estadística en 2017, siendo éste de 20,86% y partiendo de una cifra de 7,07% en 1993.

En las figuras 1 y 2 se observa que el grupo de 15-24 años (barra naranja) destaca de forma uniforme en todos los años siendo el nicho al que afecta con más frecuencia la alergia. Esta es una idea común (la alergia es una enfermedad de gente joven) pero aquí visualizamos datos cuantitativos de forma irrefutable. Como es lógico en los nichos de edad más avanzada (barras verde, azul marino y marrón) se observa un claro descenso de afectación por alergia.

La prevalencia en varones es claramente mayor en el grupo de 15-24 años con respecto a los demás grupos de edad, Desde ese grupo es como una escalera que desciende progresivamente hasta los grupos de mayor edad. En 2017 y en los años anteriores se muestra un dibujo similar excepto en 2014 en que el grupo de 25-34 años era ligeramente superior. Podemos añadir, que el segundo grupo más afectado es el de la década 25-34 años (barra gris) que además lo hace de forma homogénea en todos los años.

En mujeres no es tan evidente la preeminencia del grupo de 15-24 años, siendo superado en 2014 y 2017 por la década de 25-34 años. En general hay menos diferencias en las distintas décadas siendo la distribución más uniforme.

Resultados

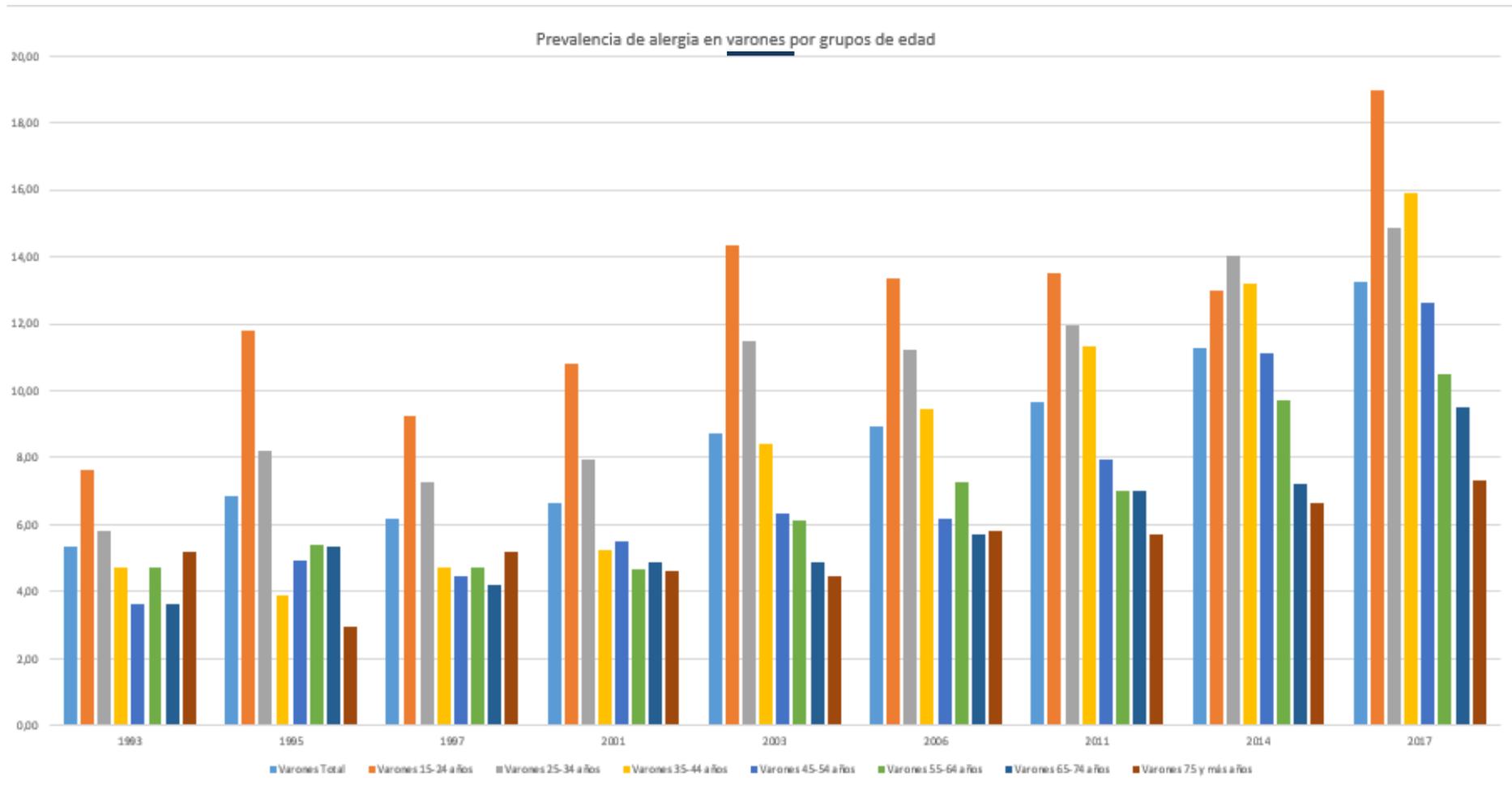


Figura 1. Diagrama de barras con el porcentaje de prevalencia en varones en ordenadas y en abscisas los diferentes grupos según el año de la encuesta con subdivisiones de colores distintos (relación de grupo de edad y sexo) con su significado en la leyenda del pie de la figura.

Resultados

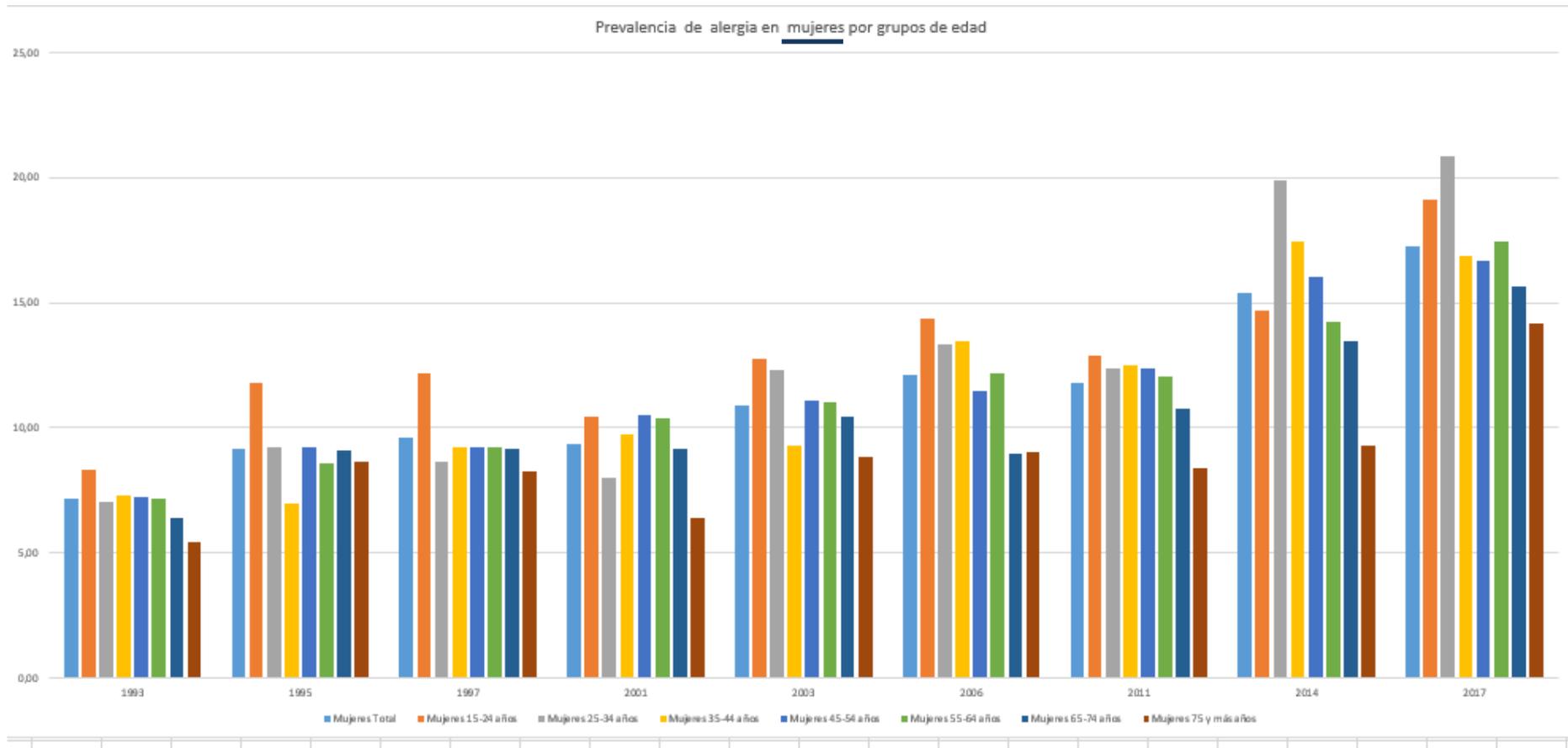


Figura 2. Diagrama de barras con el porcentaje de prevalencia en mujeres en ordenadas y en abscisas los diferentes grupos según el año de la encuesta con subdivisiones de colores distintos (relación de grupo de edad y sexo) con su significado en la leyenda del pie de la figura.

La figura 3 es una representación gráfica en forma de picos montañosos del porcentaje de prevalencia de alergia. Las cifras más bajas (en azul claro) con un porcentaje de prevalencia no superior al 5% las presentan los varones y las edades más avanzadas. El punto más alto (en azul oscuro, más del 20%) está en las mujeres entre 25-34 años. En el total de mujeres la prevalencia también es más alta y se ve que en los últimos años se eleva progresivamente hasta alcanzar cifras entre 15-20% (en amarillo). Desde 1993 las cifras ascienden con el paso de los años desde el color naranja (5-10%) pasando por el gris (10-15%) hasta alcanzar el amarillo (15-20%). Podemos observar que tanto en conjunto como por separado en ambos sexos la figura forma una escalera descendente desde las edades más tempranas hasta las más tardías.

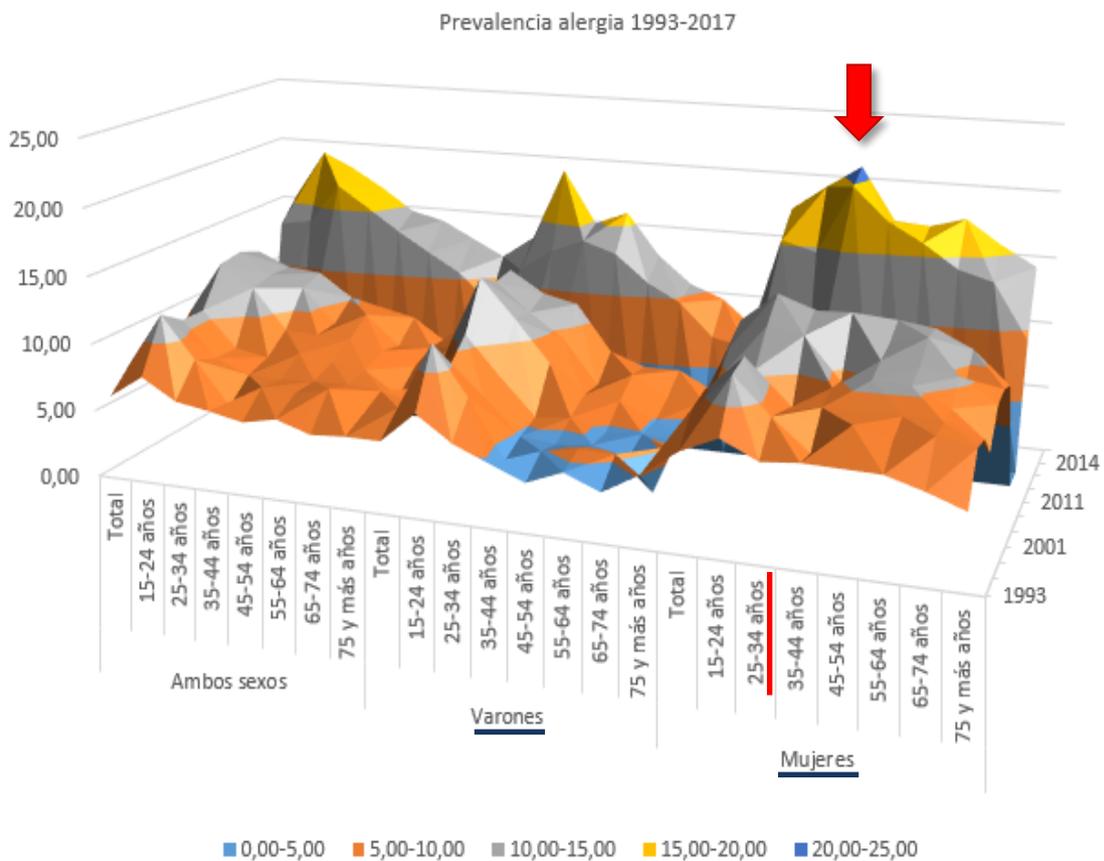


Figura 3. Representación gráfica en forma de picos montañosos del porcentaje de prevalencia de alergia diferenciándose por edad y sexo, y distinguiendo los años de medición. En la base de la figura la leyenda explica el porcentaje en forma de diferentes colores.

Resultados

La figura 4 nos muestra claramente la diferencia a favor de las mujeres en cuanto a prevalencia de alergia. La idea común de que la alergia afecta en mayor medida a las mujeres también es habitual pero en la gráfica se ve la diferencia en el porcentaje de prevalencia a favor de las mujeres con respecto a los varones, y esto se comprueba en todos los años de medición. Si comparamos en la figura 4 los colores similares observamos que el mismo color correspondiente al grupo femenino siempre está más elevado que en el masculino.

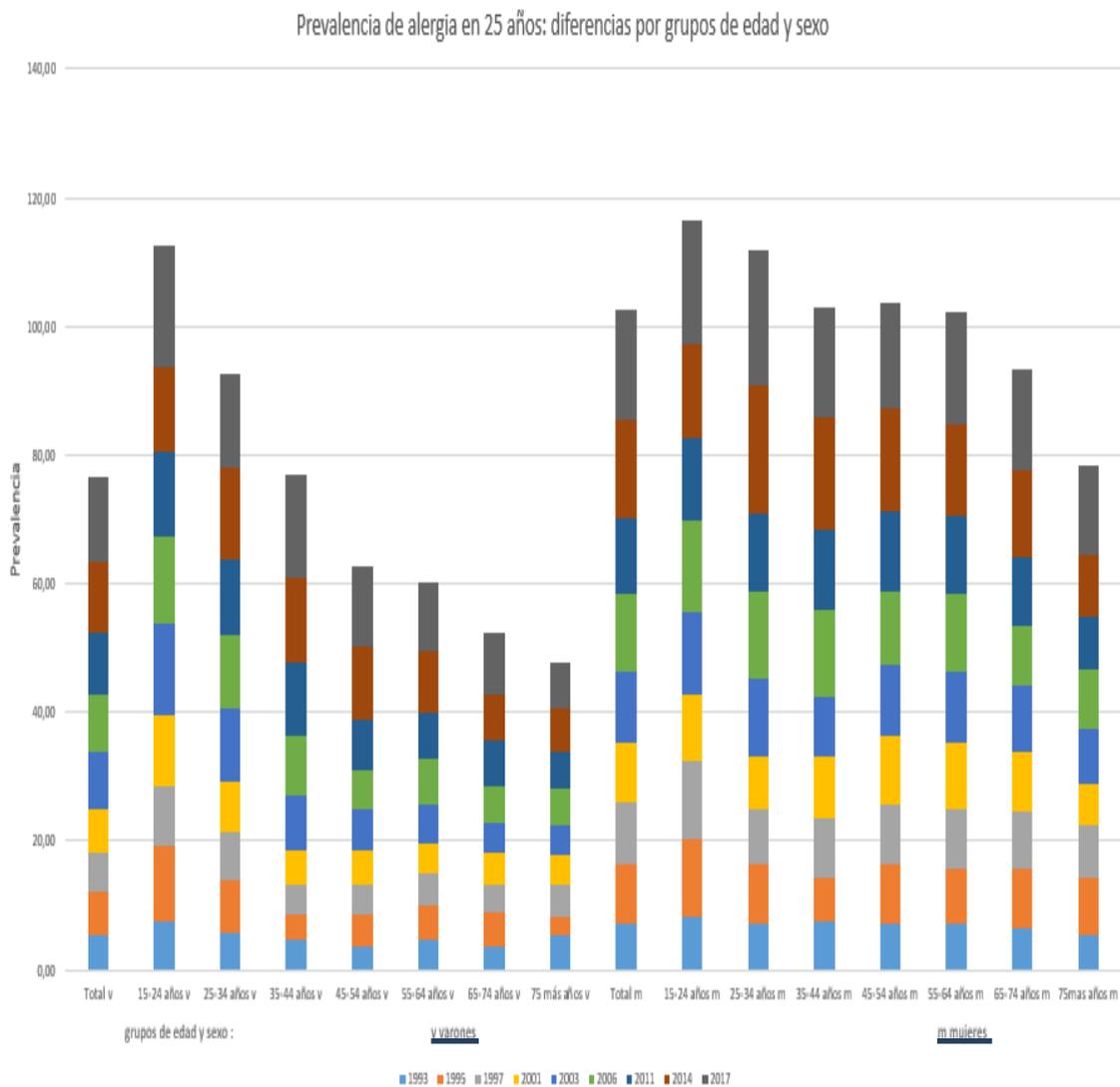


Figura 4. Diagrama de barras acumulado representando en ordenadas el porcentaje de prevalencia de alergia y en abscisas, diferenciando por sexos, la distribución por grupos de edad. Los años de medición representados por diferentes colores en la leyenda.

Resultados

Realizamos prueba de Shapiro-Wilk para verificar la normalidad de la distribución con resultado de $W= 0,9701$, $p\text{-valor}= 0,67$. Interpretación: como el $p\text{-valor}$ es mayor que $0,05$ ($0,67 > 0,05$) aceptamos la hipótesis nula (H_0), por lo que podemos afirmar que nuestros datos siguen una distribución normal.

En la figura 5 vemos una serie temporal de prevalencia de alergia desde 1993 a 2017. Lo que es más destacable es el aumento de prevalencia en 25 años desde un 6,25% a un 15,28%, el cual además es significativo estadísticamente ($p= 2,107e^{-13}$), como muestran los cálculos de la T de Student para muestras independientes tomando como referencia las cifras en 1993 y en 2017.

Vemos que el ascenso es progresivo y no se estabiliza en ninguno de los años de medición. De esta forma tan clara se visualiza y cuantifica esa idea común que se ha ido formando del aumento de la alergia con los años.



Figura 5. Serie temporal de prevalencia de alergia desde 1993 a 2017. En ordenadas se expresa el porcentaje de prevalencia y en abscisas cada uno de los años en que se efectuaron las encuestas. Numéricamente también se expresa la prevalencia sobre la gráfica y al pie como leyenda.

Resultados

Tabla 2. Cifras de estadística descriptiva de prevalencia en 25 años por grupos de edad y sexo.

	Ambos sexos								Varones								Mujeres							
	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años
Media	10,003205	12,717925	11,349222	9,9716289	9,2798445	9,0887639	8,3500394	7,3012336	8,5257	12,514033	10,297091	8,5365622	6,9647114	6,6787884	5,8241261	5,3168615	11,404469	12,942027	12,416148	11,424721	11,534134	11,356558	10,349537	8,7171396
Error estándar	0,957579	1,0250634	1,3425622	1,3308431	1,0163424	0,8670993	0,7303148	0,6071758	0,8601685	1,0865522	1,0467512	1,4241185	1,0180095	0,7264589	0,6060033	0,4271757	1,0610883	1,0049352	1,667627	1,2910049	1,0357401	1,0339208	0,9121003	0,8052565
Mediana	9,834716	13,2	11,885666	8,848842	8,7044357	8,6364918	7,6388889	7,0635792	8,7150556	12,991425	11,199925	8,4247236	6,1933166	6,1437703	5,3380783	5,2023121	10,901102	12,726904	12,302101	9,719063	11,057178	10,991763	9,1346154	8,6538462
Primer cuartil	8,0303099	10,657785	7,9774288	6,9676153	7,1764706	7,0627803	7,3514134	6,0846561	6,6279513	10,825406	7,9299768	4,7151277	4,9261084	4,7111882	4,8842238	4,6154764	9,3423137	11,804384	8,6601307	9,2156863	9,2342342	9,2077088	9,1116173	8,2568807
Tercer cuartil	10,74	13,802694	12,23917	11,92	10,18	9,788639	9,04	7,7019181	9,66	13,52	11,93	11,35	7,94	7,251579	7,03	5,8180628	12,100573	14,337	13,355604	13,453887	12,39	12,178722	10,78	9,0397544
Varianza	8,2526174	9,4567938	16,22226	15,940289	9,2965661	6,76675	4,8002368	3,3179624	6,6590079	10,625361	9,8611923	18,253021	9,3270894	4,7496824	3,3051597	1,6423118	10,133176	9,0890526	25,028817	15,000244	9,6548182	9,6209308	7,487342	5,8359423
Desviación estándar	2,8727369	3,0751901	4,0276867	3,9925292	3,0490271	2,6012978	2,1909443	1,8215275	2,5805054	3,2596565	3,1402535	4,2723554	3,0540284	2,1793766	1,8180098	1,2815272	3,1832649	3,0148056	5,0028809	3,8730148	3,1072203	3,1017625	2,7363008	2,4157695
Curtosis	-0,0536695	1,9226484	-0,8175273	-1,1240611	-0,1639732	0,1587981	1,2499445	3,1867897	-0,3184022	1,2371062	-1,336247	-0,9920896	0,0349054	-0,4366306	0,9138131	0,6211766	0,0594314	1,844972	-0,4310385	-1,0815722	-0,3119697	0,6522908	0,85037	3,4908961
Asimetría	0,7347247	0,7187737	0,582947	0,564006	0,8604573	0,8844799	0,9019368	1,5271332	0,6410809	0,5511782	0,0556016	0,5595576	1,0546064	0,9099483	1,0209345	-0,2893948	0,7796408	0,7620118	0,8925017	0,575312	0,6471994	0,7664291	0,8577284	1,2904767
Intervalo	9,0316542	11,105004	11,449487	10,900329	9,1747585	8,060477	7,5757138	6,1007403	7,910202	11,358063	9,0444526	11,987022	8,9721688	5,8218824	5,8802689	4,3911715	10,07655	10,804695	13,785771	10,451892	9,3998241	10,228487	9,2458404	8,7272987
Mínimo	6,2455575	7,9459682	6,409997	5,465587	5,4628491	5,9836876	5,1623917	5,3373844	5,3218177	7,6101884	5,7933364	3,8934426	3,6439835	4,6877942	3,638897	2,9411765	7,1411001	8,3330661	7,0725021	7	7,2648289	7,1935537	6,3993958	5,4341701
Máximo	15,277212	19,050973	17,859484	16,365916	14,637608	14,044165	12,738106	11,438125	13,23202	18,968252	14,837789	15,880465	12,616152	10,509677	9,5191659	7,332348	17,21765	19,137761	20,858273	17,451892	16,664653	17,42204	15,645236	14,161469
Suma	90,028845	114,46133	102,143	89,74466	83,518601	81,798875	75,150355	65,711103	76,7313	112,6263	92,673823	76,82906	62,682403	60,109096	52,417135	47,851754	102,64022	116,47825	111,74533	102,82249	103,80721	102,20902	93,14583	78,454257

Como observamos en la tabla 2, de la estadística descriptiva de cifras de prevalencia de alergia podemos obtener datos. En cifras totales de ambos sexos destaca la media del 12,71% del nicho de 15-24 años, siendo en varones de 12,51% y en mujeres de 12,94%. En el resto de grupos cifras abrumadoramente superiores en mujeres, en algunos casos como en la década de 45-54 años pudiendo ser de casi el doble (6,96% en hombres y 11,53% en mujeres). Hay una disminución progresiva con la edad, con más de 75 años la cifra en mujeres disminuye de un 11,40% (cifra total en mujeres) hasta un 8,41%. En los grupos de 15-24 años y 25-34 años, las cifras son similares en hombres y mujeres, pero después las cifras de alergia crecen a expensas sobre todo de mujeres.

Como observamos en la tabla, todas las medianas se mantiene en valores cercanos a la media, lo cual habla de una distribución centrada. Esto se confirma viendo las cifras de curtosis (mínima en el caso de la cifra total con un valor de -0,05), con variaciones positivas o negativas según el grupo de edad (en algunos casos tan mínimas como en el grupo de varones 45-54 años de 0,03). Lo mismo podemos decir del coeficiente de asimetría (destaca el grupo de varones de 25-34 años con 0,05), aunque en el caso de este parámetro existe una homogeneidad al situarse los valores hacia la derecha.

En cuanto a la dispersión de la muestra, en la cifra total se establece entre un 20%, observando el valor de la desviación y la varianza. En el resto de los grupos de edad también entre un 20-30% (aunque en algunos casos hay valores entre el 40-50%).

La figura 6 es un biplot y nos muestra la dificultad para visualizar y entender múltiples datos. Es fácil hacerse una representación hasta de 3 variables pero en cuanto aumentamos el número nos cuesta situarlas en el espacio para poder interpretarlas gráficamente. La tabla 3 y la figura 6 es un intento de superar esta dificultad. La tabla 3 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 6 y que corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de observaciones de edad y sexo. La proximidad entre puntos se interpreta en términos

Resultados

de similitud, el ángulo en términos de correlación (si los vectores son perpendiculares $\cos=0$ las variables son independientes, si el $\cos=180^\circ$ hay una correlación positiva alta). Vemos que los años 2014 y 2017 se sitúan en posición ortogonal con respecto a los primeros años de medición, indicando la diferencia de relación entre ellos ya que las cifras de prevalencia son distintas.

Como ejemplo, vemos que se producen agrupaciones de forma natural como en los grupos 2, 10, 18 que corresponden todos ellos al grupo de edad de 15-24 años tanto en sus cifras totales como en hombres y mujeres, y sus proyecciones sobre los ejes nos dan valores de prevalencia altos. Así mismo, el nicho 19 (mujeres entre 25-34 años) se proyecta en el punto máximo de prevalencia en 2017. Por el contrario, los grupos de menor prevalencia, como es el caso del nicho 5 (ambos sexos de 45-54 años) y los que recogen las edades más avanzadas (15, 16, 23, 24), se proyectan en zonas de baja prevalencia.

Tabla 3. Clave numérica para identificar cada una de las variables que aparecen en la figura 6 y que corresponden a diferentes nichos de edad y sexo.

	1	2	3	4	5	6	7	8
	<u>Ambos sexos</u>							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	9	10	11	12	13	14	15	16
	<u>Varones</u>							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	17	18	19	20	21	22	23	24
	<u>Mujeres</u>							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	

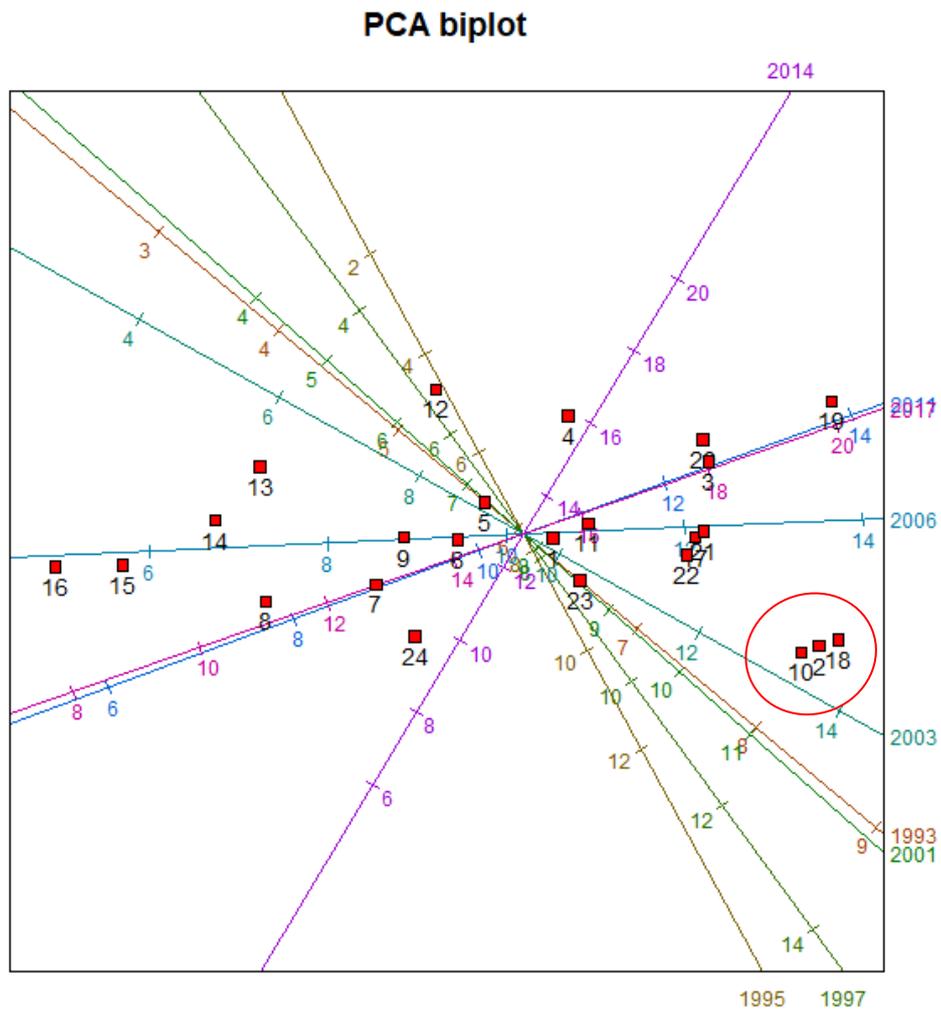


Figura 6. Biplot representando una serie de ejes que son los años de medición en que se efectúan las observaciones, incorporando en ellos las cifras de prevalencia. Proyectándose sobre los ejes cada una de las observaciones numeradas que son los grupos de edad y sexo.

Resultados

4.1.2. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años con respecto al tipo de población

Si tenemos en cuenta no solo la edad y el sexo sino también el tipo de población en cuanto al número de habitantes (menos de 10.000 habitantes, entre 10.000 y 50.000 habitantes, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes), podemos observar la evolución de prevalencia de alergia como muestra la tabla 4.

En 2017 hay más alergia en poblaciones de más de 400.000 habitantes tanto en hombres (15,06%) como en mujeres (22,06%). En 2014 también, pero es menos acusado. En 2011 el máximo está en poblaciones de 100.000 a 400.000 habitantes (en total de ambos sexos 11,83% y en mujeres 14,42%) pero de modo menor. En 2006 no había grandes diferencias por el tipo de población. Las cifras más bajas corresponden a poblaciones de 10.000 habitantes y de 10.000-50.000. Excepto en 2003, 2006 y 2011, el resto de los años tiene su mayor cifra en poblaciones de más de 400.000 habitantes. Tanto en varones como en mujeres en 2014 y en 2017 es donde las cifras son más elevadas en poblaciones de más de 400.000 habitantes.

Tabla 4. Evolución de la prevalencia de alergia desde 1993 a 2017 utilizando como variables los tipos de población según números de habitantes y ofreciendo 3 divisiones: cifras totales, varones y mujeres.

	Total																	
	Ambos sexos					Varones					Mujeres							
	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
1993	6,25	5,00	5,47	6,84	6,58	9,33	5,32	4,32	4,64	6,05	5,69	7,52	7,14	5,67	6,31	7,61	7,42	10,99
1995	8,05	6,46	6,17	9,00	9,18	10,53	6,83	7,20	4,72	7,72	7,21	8,09	9,19	5,80	7,56	10,19	10,98	12,82
1997	7,95	6,19	7,19	8,20	7,91	11,02	6,18	4,86	6,06	3,15	6,01	9,56	9,60	7,44	8,29	13,18	9,63	12,36
2001	8,03	7,06	8,33	8,82	8,03	8,52	6,63	5,10	7,54	8,19	6,71	6,51	9,34	8,90	9,08	9,40	9,26	10,39
2003	9,83	10,27	9,40	10,65	10,02	9,13	8,72	8,82	8,78	9,65	8,43	8,28	10,90	11,64	10,00	11,65	11,56	9,88
2006	10,54	10,06	10,44	10,54	11,04	10,59	8,91	8,24	8,21	9,10	9,94	9,21	12,10	11,72	12,60	11,89	12,18	11,91
2011	10,74	9,91	10,51	10,53	11,83	10,87	9,66	10,69	8,49	10,05	9,07	10,63	11,77	9,10	12,41	11,01	14,42	11,08
2014	13,37	12,22	12,99	12,74	13,89	15,06	11,25	10,58	11,32	9,06	12,23	12,43	15,38	13,92	14,61	16,33	15,38	17,38
2017	15,28	12,95	15,18	13,36	15,93	18,66	13,23	11,71	13,24	12,03	13,95	15,06	17,22	14,22	17,00	14,63	17,71	22,06

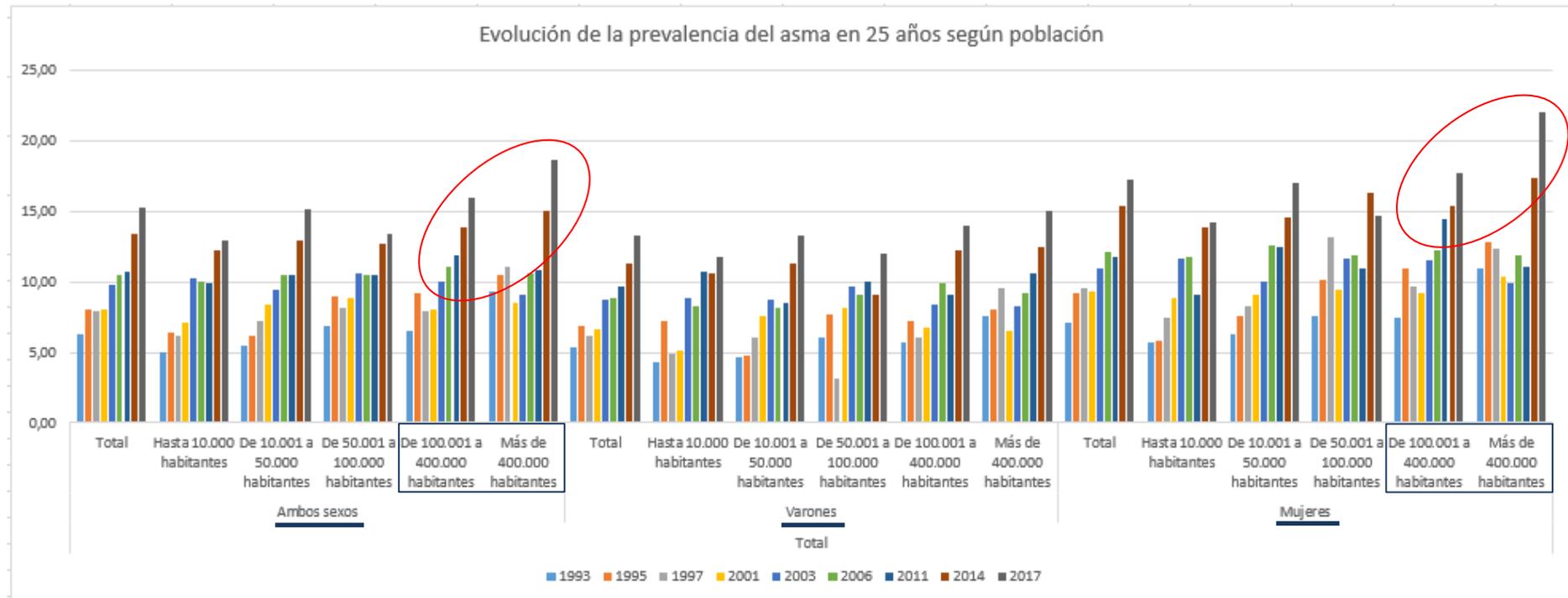


Figura 7. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de población según número de habitantes con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Resultados

La figura 7 muestra de manera clara la mayor prevalencia en mujeres y también el crecimiento de la alergia en todos los grupos de población. En las poblaciones de más de 400.000 habitantes el crecimiento de la alergia en los últimos registros ha subido claramente, siendo en mujeres más marcado, mientras que en la década de 2003 a 2011 (colores azules oscuro y verde) se mantenía constante independientemente del número de habitantes.

Por poblaciones vemos que en los últimos años ha aumentado sobre todo la prevalencia en poblaciones de 100.000 a 400.000 habitantes y en poblaciones de más de 400.000 habitantes. Mientras que en el resto de poblaciones más pequeñas las cifras son más similares y con menos diferencias entre ellas. Esto es más acusado en la población femenina.

La diferencia de prevalencia en poblaciones de más de 400.000 habitantes en 25 años (9,33% en 1993 respecto al máximo de 18,66% en 2017) nos muestra un crecimiento del 50% que sería algo inferior al crecimiento de la alergia en general (de un 59,09%) sin tener en cuenta el tipo de población (6,25% en 1993 respecto al máximo de 15,28% en 2017).

Tabla 5. Estadística descriptiva de la evolución de la alergia en 25 años con respecto al tipo de población en el que vive el alérgico.

	Total																	
	Ambos sexos						Varones					Mujeres						
	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
Media	10,003607	8,9009499	9,5204866	10,077902	10,491226	11,523579	8,5262528	7,9467471	8,1113983	8,3338025	8,8042742	9,7006979	11,404664	9,8237919	10,874988	11,764993	12,061609	13,206525
Error típico	0,9576177	0,9394618	1,0567661	0,7006015	1,0086394	1,0906258	0,8602597	0,919239	0,9518787	0,8494489	0,9440864	0,8887312	1,0610968	1,076544	1,1749502	0,8900593	1,0881414	1,3264297
Mediana	9,834716	9,9059392	9,3974145	10,533684	10,020639	10,592024	8,7150556	8,2365064	8,2101613	9,0569093	8,4287417	9,2125562	10,901102	9,1036776	10,002958	11,646637	11,564701	11,909625
Desviación estándar	2,8728532	2,8183854	3,1702984	2,1018045	3,0259063	3,2718773	2,5807792	2,757717	2,8556361	2,5483467	2,8322592	2,6661935	3,1832903	3,2296319	3,5248505	2,6701779	3,2644242	3,979289
Varianza de la muestra	8,2532852	7,9432965	10,050792	4,417582	9,1561092	10,705181	6,6604214	7,6050033	8,1546575	6,4940711	8,0216919	7,1085879	10,133337	10,430522	12,424571	7,1298501	10,656465	15,834741
Curiosis	-0,055063	-1,477866	-0,327931	-0,557558	-0,314754	2,1955525	-0,320945	-1,623765	-0,178061	1,4541268	-0,311558	0,8419507	0,0589653	-1,443138	-0,69705	-0,128197	-0,481259	2,550227
Coefficiente de asimetría	0,734136	0,0566038	0,5365869	0,1917681	0,6316584	1,6272048	0,640231	-0,058135	0,5417259	-0,901524	0,798372	1,0658181	0,7793885	0,071409	0,4985	0,2717183	0,4271809	1,7413807
Rango	9,0316542	7,9503484	9,7145818	6,519078	9,3520053	10,147558	7,910202	7,3925278	8,5932232	8,880164	8,2622503	8,5535063	10,07655	8,5499452	10,697287	8,725352	10,291927	12,177505
Mínimo	6,2455575	4,9995782	5,467226	6,8443486	6,5815848	8,5154608	5,3218177	4,3183315	4,6429659	3,1496063	5,6925376	6,5113362	7,1411001	5,6700945	6,3071867	7,6073494	7,4195523	9,8776219
Máximo	15,277212	12,949927	15,181808	13,369427	15,93359	18,663019	13,23202	11,710859	13,236189	12,02977	13,954788	15,064843	17,21765	14,22004	17,004474	16,332701	17,711479	22,055127
Suma	90,032463	80,108549	85,684379	90,701122	94,421031	103,71222	76,736275	71,520724	73,002585	75,004222	79,238468	87,306281	102,64198	88,414127	97,87489	105,88437	108,55448	118,85873

La tabla 5 complementa toda la estadística descriptiva de la evolución de la alergia en 25 años con respecto al tipo de población en el que vive el alérgico. En poblaciones de más de 400.000 habitantes de media en 25 años es el 11,52% y ha ido subiendo progresivamente en 25 años. Valores máximos en poblaciones de más de 400.000 habitantes, totales de 11,52%, en mujeres 13,20% y en varones 9,70%. Si tenemos en cuenta la diferencia de prevalencia en 2017 entre las poblaciones de más de 400.000 habitantes y la cifra total media para todas las poblaciones (18,66% frente a 15,28%) vemos que es de un 3,38 que representa un 18,11% a favor de la primera.

La distribución es muy centrada con medianas de valores muy próximos a la media. La curtosis es, sobre todo, moderadamente negativa con desviación izquierda. El coeficiente de asimetría es ligeramente positivo en su mayor parte.

La dispersión de la muestra según la desviación típica es de alrededor de un 20% en la mayor parte de grupos.

El test de normalidad de Shapiro-Wilk nos muestra un valor de $W = 0,88542$, p -valor = 0,8601, por lo que la distribución es normal.

El aumento de alergia en todo tipo de poblaciones ha sido significativo estadísticamente $p < 2,2 \cdot 10^{-16}$ al aplicar la T de Student a los grupos inicial (1993) y final (2017). Así mismo, si comparamos solo las poblaciones de más de 400.000 habitantes en lo que respecta a las medias de los grupos total, mujeres y hombres, vemos que el aumento en 25 años también es significativo con $p = 1,687 \cdot 10^{-12}$.

Resultados

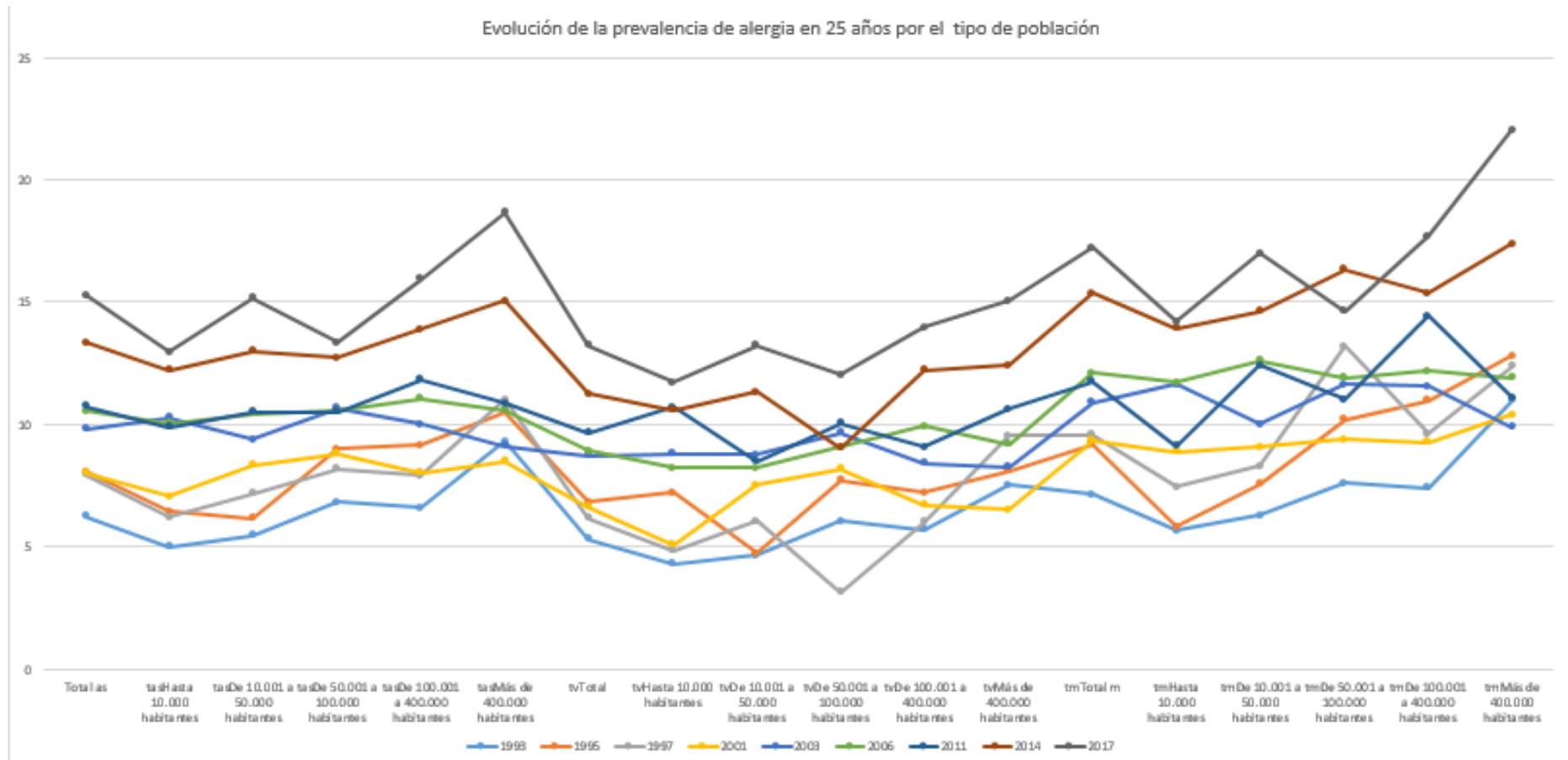


Figura 8. Serie temporal de la evolución de prevalencia de alergia en 25 años (en la leyenda a pie de tabla se explica el color de cada año). En abcisas el tipo de población y con 3 divisiones: as –cifra total ambos sexos-, v–varones-, m–mujeres-. En ordenadas el porcentaje de prevalencia.

La figura 8 nos da una perspectiva de todos los nichos tabulados en la tabla 5 de estadística descriptiva diferenciando además con el dibujo de cada uno de los años de medición en colores distintos. Es una serie temporal de la evolución de prevalencia de alergia en 25 años (en la leyenda a pie de tabla se explica el color de cada año). En abcisas el tipo de población y con 3 divisiones: as –cifra total ambos sexos-, v –varones, m –mujeres-. En ordenadas se muestra el porcentaje de prevalencia. Nos permite visualizar el ascenso marcado en los últimos años sobre todo en las grandes poblaciones, y que se da tanto en cifras totales como por separado en ambos sexos.

Vemos en abcisas el tipo de población y en ordenadas el porcentaje según los años. Las poblaciones de más de 400.000 y de 100.000-400.000 habitantes están en los valores más altos sobre todo en 2014 y 2017, siendo esto particularmente mayor en mujeres.

En poblaciones de más de 400.000 habitantes de media en 25 años es el 11,52% y ha ido subiendo progresivamente en 25 años. En los últimos años ha aumentado sobre todo la prevalencia en poblaciones de 100.000 a 400.000 habitantes y en poblaciones de más de 400.000 habitantes. Mientras que en el resto de poblaciones más pequeñas las cifras son más similares y con menos diferencias entre ellas.

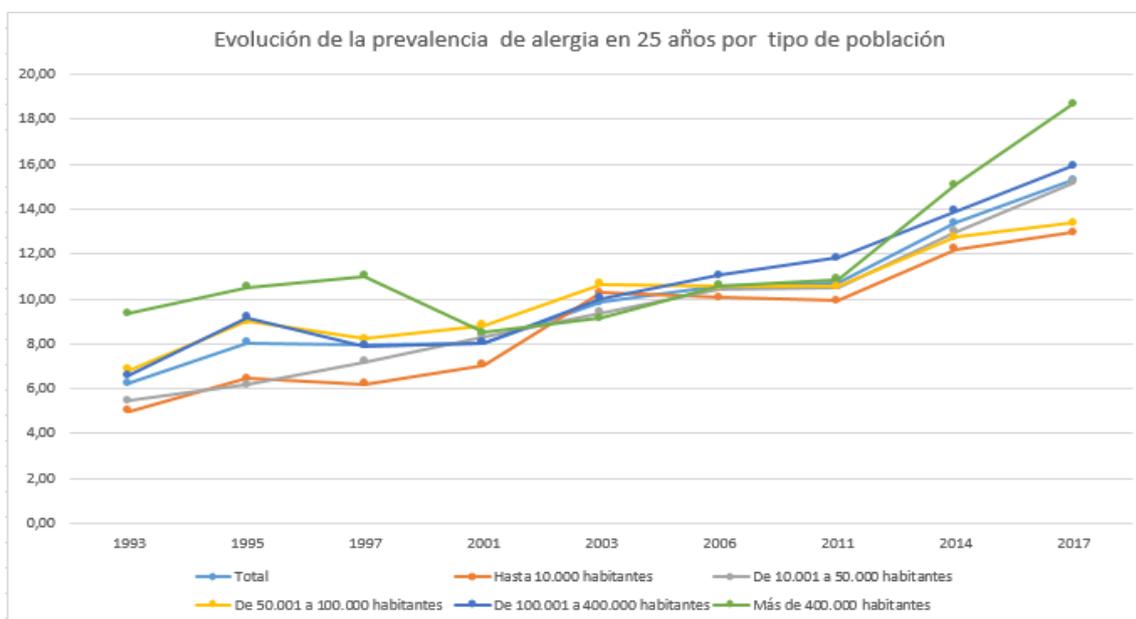


Figura 9. Serie temporal de 25 años representando en abcisas las series de años y en ordenadas el porcentaje de prevalencia de alergia en cifras totales, caracterizando por colores el tipo de población.

La figura 9 nos muestra en abcisas la serie de años y en ordenadas el porcentaje de prevalencia en cifras totales, apareciendo además en diferentes colores cada uno de los tipos de población. Nos señala que las poblaciones de más de 400.000 habitantes están en los valores más altos. Gráficamente se ve el aumento de la alergia con los sucesivos años que hasta los últimos años no se refleja en el tipo de población, pero que en las últimas mediciones se decanta el incremento en las poblaciones de más de 400.000 habitantes (en verde) y a continuación las de 100.000 a 400.000 habitantes, ocupando el lugar más bajo las que no alcanzan 10.000 habitantes (en naranja).

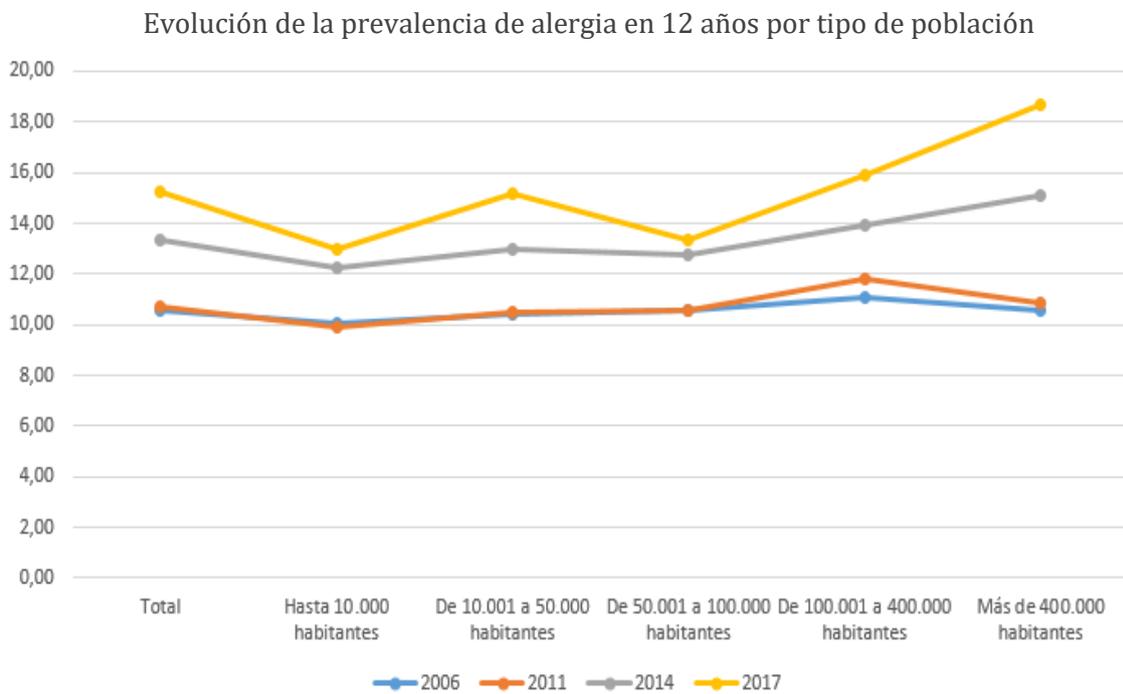


Figura 10. Serie temporal de la evolución de prevalencia de alergia en sucesivos años (en la leyenda a pie de tabla se explica el color que corresponde a cada año) teniendo en abcisas el tipo de población y en ordenadas el porcentaje de prevalencia.

La figura 10 nos da una serie temporal de la evolución de prevalencia de alergia en sucesivos años de una serie más corta (en la leyenda a pie de tabla se explica el color que corresponde a cada año) teniendo en abcisas el tipo de población y en ordenadas

Las variables que aparecen en la tabla 6 son las que se presentan en el biplot y corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición (numerados por orden, desde el 1 al 9, desde 1993 a 2017), situándose sobre los ejes los porcentajes de prevalencia y proyectándose en los ejes cada uno de los nichos población y sexo.

La proximidad entre puntos se interpreta en términos de similitud, el ángulo en términos de correlación (si los vectores son perpendiculares $\cos=0$ las variables son independientes, si el $\cos=180$ hay una correlación positiva alta) y las proyecciones de los años sobre las variables permiten obtener los valores de la muestra de datos.

Las cifras más altas de prevalencia se presentan en los últimos años como se puede apreciar ya que los ejes 7, 8 y 9 corresponden a 2011, 2014 y 2017 mostrando superposición de líneas y por tanto ángulos cercanos a 180° , es decir con una alta correlación. Los nichos 24, 30, 36 y 72 que corresponden a los habitantes de poblaciones de más de 400.000 habitantes, tanto de ambos sexos como por separado varones y mujeres, se sitúan en la zona periférica del gráfico con las proyecciones más elevadas de prevalencia. Esto mismo ocurre con los nichos 29, 23 y 35 pertenecientes a ciudades entre 100.000-400.000 habitantes. En zonas periféricas pero proyectándose en los ejes en los valores menores de prevalencia, tenemos a los nichos 40, 46, 52, 64, 76, 88, todos ellos representando a poblaciones entre 50.000-100.000 habitantes.

Resultados

Tabla 6. Clave numérica para identificar cada una de las variables que aparecen en la figura 11 y que corresponden a diferentes nichos de edad y sexo.

Total																	
Ambos sexos					Varones					Mujeres							
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
15-24 años																	
Ambos sexos					Varones					Mujeres							
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
25-44 años																	
Ambos sexos					Varones					Mujeres							
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
45-64 años																	
Ambos sexos					Varones					Mujeres							
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
65 y más años																	
Ambos sexos					Varones					Mujeres							
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes

4.1.3. Estadística descriptiva de la prevalencia de alergia según nivel de estudios

Si tenemos en cuenta los niveles de estudio considerando como tales: sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado; podemos ver en la tabla 7 que los valores de alergia son más altos en personas con nivel superior tanto en cifras totales como si separamos en hombres y mujeres, ocurriendo este hecho independientemente del año en que se analice.

La diferencia en 12 años entre los datos, con la aplicación del test de Student es significativa estadísticamente $p = 2,739 \times 10^{-16}$.

La distribución sigue un patrón normal (Shapiro-Wilk $W = 0,98798$, $p\text{-valor} = 0,5835$).

Tabla 7. Prevalencia de alergia en 12 años diferenciando grupos de niveles de estudios junto con sexo.

	Total																	
	Ambos sexos						Varones					Mujeres						
	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
2006	10,54	9,97	9,21	11,16	10,38	12,35	8,91	7,01	7,57	8,12	8,76	11,87	12,10	12,42	10,69	14,15	12,03	12,87
2011	10,74	8,50	9,49	9,59	12,04	13,58	9,66	7,23	7,49	8,61	10,54	13,72	11,77	9,61	11,25	10,57	13,50	13,45
2014	13,37	11,18	11,12	13,70	14,57	15,14	11,25	8,43	8,42	11,71	12,96	12,98	15,38	13,35	13,46	15,90	16,19	17,18
2017	15,28	13,72	13,69	13,88	16,21	17,77	13,23	8,94	11,32	11,98	14,64	16,25	17,22	17,17	15,75	16,10	17,73	19,23

En la figura 12 tenemos las cifras numéricas de la tabla 7 en forma de diagrama de barras y así podemos visualizar esta misma circunstancia. A su vez, las cifras más bajas de prevalencia coinciden con el grupo de personas sin estudios. Con respecto al nivel de tercer grado, en cifras totales en ambos sexos se alcanzan valores que aumentan desde 12,35% a 17,77% en el último año de medición, siendo éste el punto máximo. En cambio, en las personas sin estudios oscila entre 9,97% y 13,72%. En mujeres con estudios de tercer grado es más marcado y se alcanza el punto máximo de toda la estadística que es de 19,23% en 2017.

Resultados

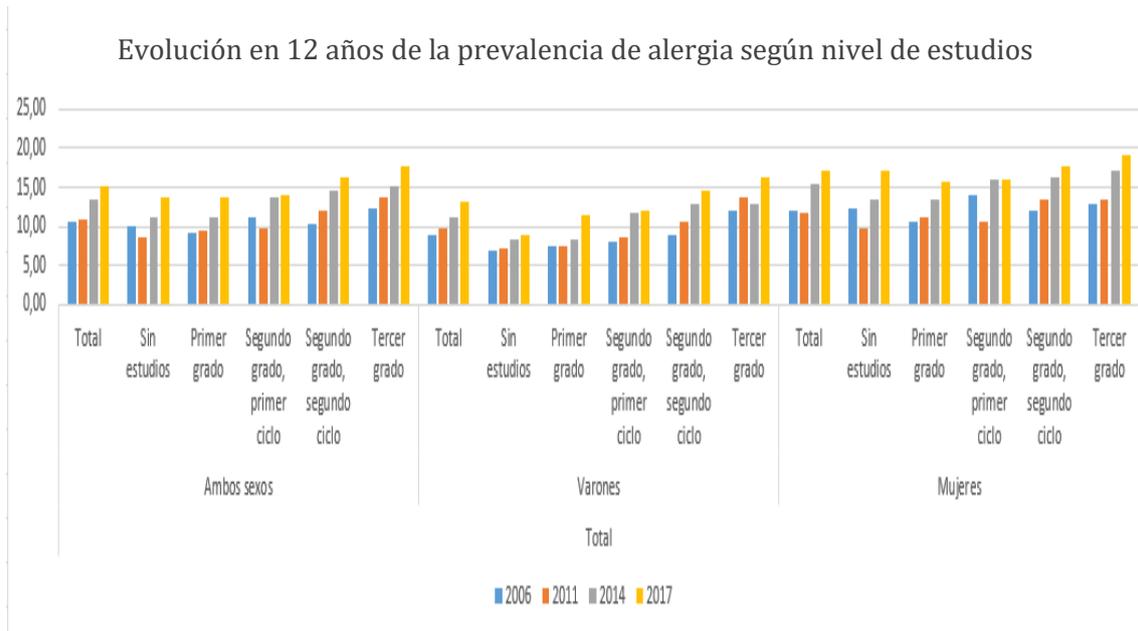


Figura 12. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de niveles de estudios con 3 divisiones: cifras totales en ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

En la figura 12 podemos visualizar que las cifras más bajas de prevalencia coinciden con el grupo de personas sin estudios y que las más altas son para los niveles superiores de tercer grado. De forma que tanto en cifras totales como para varones y en menos grado en mujeres se forma una escalera ascendente de prevalencia de alergia hacia los niveles más altos de estudios.

El dibujo que vemos en la figura 13 que forma la prevalencia de alergia (obtenida con el promedio de la evolución en 12 años) con el nivel de estudios, es como una escalera, conforme aumenta el nivel de estudios aumenta la alergia. En el caso de las mujeres, la pendiente es menor pero claramente vemos la diferencia entre el grupo sin estudios y el grupo de más nivel.

Resultados

El test de Student reveló significatividad en la diferencia de cifras entre la alergia del nivel sin estudios y la prevalencia con estudios superiores, $t = 2,8577$, $p\text{-valor} = 0,01305$. Se aplicó dicho test después de comprobar la homogeneidad de las varianzas con el test de Barlett ($F = 1,4831$, $p\text{-valor} = 0,6159$), y no rechazar la hipótesis nula de homocedasticidad.

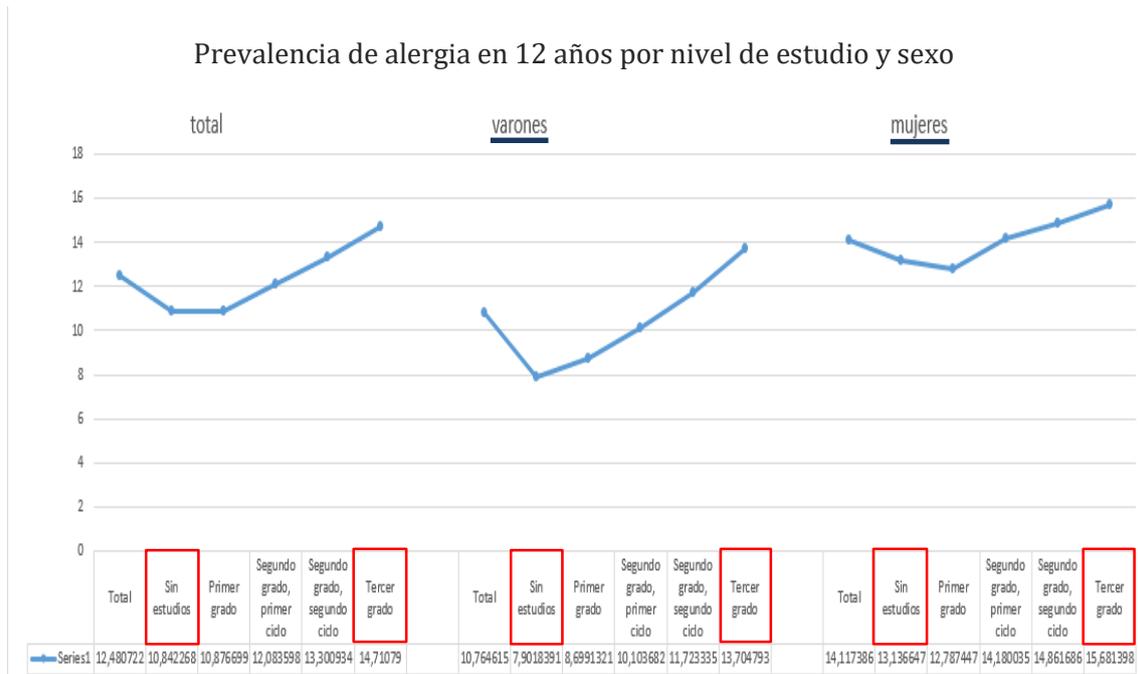


Figura 13. Promedio de la prevalencia de alergia en 12 años, expresando en ordenadas el porcentaje de prevalencia y en abscisas los distintos grupos en función del nivel de estudios y diferenciando en cifras totales y sexo.

Resultados

Tabla 8. Estadística descriptiva de la evolución de la prevalencia de alergia en 12 años con respecto a nivel de estudios y sexo.

	Ambos sexos																	
						Varones					Mujeres							
	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
Media	12,480722	10,842268	10,876699	12,083598	13,300934	14,71079	10,764615	7,9018391	8,6991321	10,103682	11,723335	13,704793	14,117386	13,136647	12,787447	14,180035	14,861686	15,681398
Error típico	1,1327733	1,1046594	1,0284439	1,0361715	1,2964257	1,169114	0,9560397	0,4658405	0,8977739	1,0104923	1,2999659	0,9282215	1,3157447	1,5609618	1,1533297	1,2800581	1,2873436	1,5197983
Mediana	12,054348	10,574906	10,304121	12,430665	13,305321	14,362397	10,457603	7,8290821	7,9919512	10,157429	11,747357	13,348921	13,740066	12,881839	12,355113	15,024842	14,844533	15,313556
Desviación estándar	2,2655466	2,2093188	2,0568877	2,0723431	2,5928514	2,3382279	1,9120793	0,9316809	1,7954779	2,0209845	2,5999317	1,8564429	2,6314894	3,1219237	2,3066595	2,5601162	2,5746872	3,0395967
Varianza de la muestra	5,1327014	4,8810896	4,2307871	4,2946057	6,7228782	5,4673097	3,6560474	0,8680294	3,223741	4,0843785	6,7596451	3,4463803	6,9247363	9,7464076	5,320678	6,554195	6,6290141	9,2391479
Curtosis	-2,7262193	0,405467	0,3184584	-3,2991273	-2,4979517	-0,070551	-0,9359278	-4,3249119	2,7891712	-5,6294156	-2,2153588	1,4353677	-3,6472352	1,198698	-1,5130607	1,3780567	-2,9701542	-3,7086789
Coefficiente de asimetría	0,571917	0,640232	1,1473363	-0,4671448	-0,0069355	0,7348675	0,7095373	0,2048199	1,6969978	-0,0340903	-0,0393633	1,0319582	0,3914746	0,4708939	0,7285256	-1,38213	0,0253998	0,3474129
Rango	4,7402309	5,221976	4,483281	4,2885332	5,8266769	5,4220763	4,3207866	1,9290529	3,8230386	3,8547839	5,8885973	4,371386	5,4458888	7,5561243	5,0551937	5,5257286	5,7028275	6,3590887
Mínimo	10,536981	8,4986419	9,2076359	9,5922634	10,38321	12,348145	8,911233	7,0100697	7,4947937	8,1225427	8,7550139	11,874973	11,771762	9,6133937	10,692184	10,572364	12,027426	12,869696
Máximo	15,277212	13,720618	13,690917	13,880797	16,209887	17,770221	13,23202	8,9391226	11,317832	11,977327	14,643611	16,246359	17,21765	17,169518	15,747378	16,098092	17,730253	19,228784
Suma	49,922889	43,369073	43,506795	48,334391	53,203738	58,843161	43,058459	31,607357	34,796528	40,414728	46,89334	54,819174	56,469544	52,546589	51,149789	56,72014	59,446745	62,725592

La tabla 8 nos revela la estadística descriptiva de la evolución de la prevalencia de alergia en 12 años con respecto al nivel de estudios y sexo.

La curtosis y la asimetría son variables según el nivel de estudios, cambiando la forma de distribución de la muestra según el citado nivel.

En cuanto a la media, las cifras más altas corresponden al grupo de más nivel (3º grado): 14,71% en total, 13,70% en varones y 15,68% en mujeres.

La dispersión de la muestra se mantiene entre un 10-20% como deducimos de las cifras de desviación estándar y varianza.

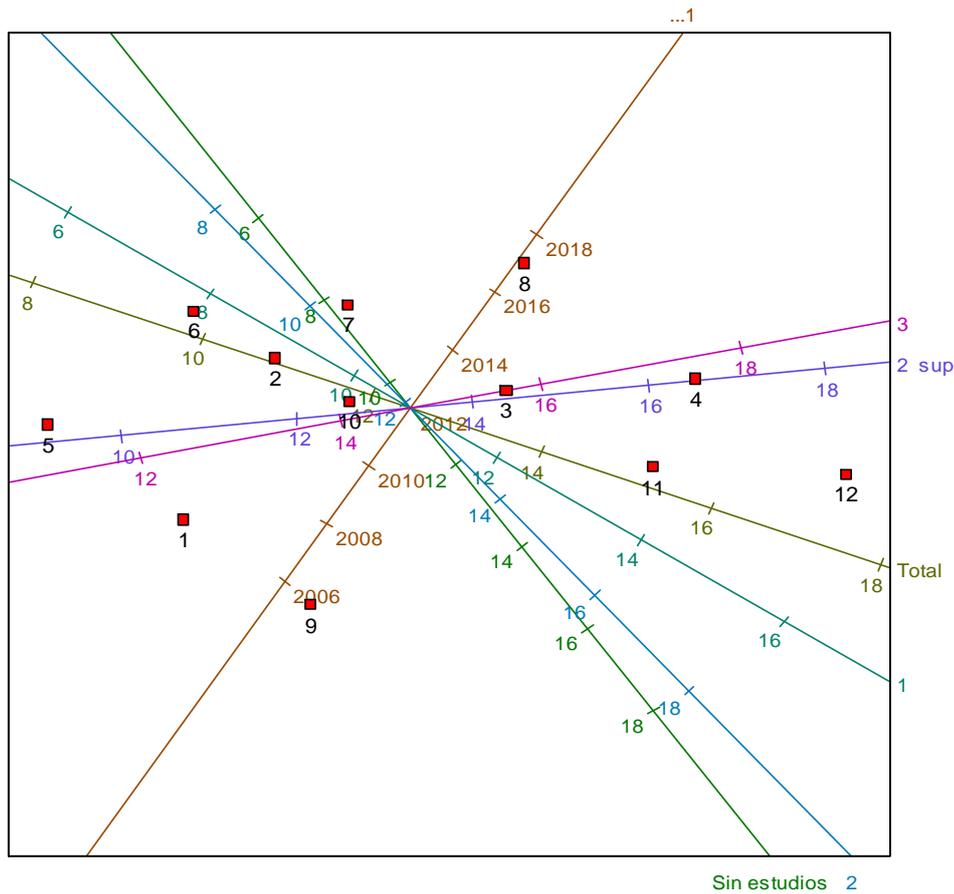


Figura 14. Biplot de prevalencia de alergia según estudios representando en los ejes los niveles de estudios (sin estudios, 1º, 2º, 2º superior, 3º) y en eje diagonal los años de medición. La variables representan los grupos por sexo y nivel de estudios.

Resultados

Tabla 9. Clave para entender la figura 14 con las variables numeradas en razón de sexo y nivel de estudios.

	1	2	3	4	5	6	7	8	9	10	11	12
	Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo	Tercer grado	

Para la representación del biplot hemos utilizado las variables que aparecen en la tabla 9 y que corresponden a los grupos por sexo y niveles de estudios. Cada uno de los ejes corresponde a los distintos niveles de estudios (numerados por orden desde primarios 1º a superiores 3º), situándose sobre los ejes los porcentajes de prevalencia.

La proximidad entre puntos se interpreta en términos de similitud, el ángulo en términos de correlación (si los vectores son perpendiculares $\cos=0$ las variables son independientes, si el $\cos=180$ hay una correlación positiva alta) y las proyecciones de los años sobre las variables permiten obtener los valores de la muestra de datos.

Nula correlación entre sin estudios, estudios primarios y 3º como vemos por la posición ortogonal de los ejes, siendo muy alta entre 3º y 2º superior al notar la superposición de estos últimos ejes con un ángulo próximo a 180º; 2º primer ciclo y sin estudios muy cercanos. En las proyecciones de mayor prevalencia sobre los ejes se sitúan observaciones que recogen estudios superiores como los nichos 11 y 12.

4.1.4. Estadística descriptiva de la prevalencia de la alergia según nivel económico

Para caracterizar el nivel socio-económico hemos usado la clase social ocupacional, la cual obtiene la clase social del individuo que trabaja o ha trabajado y del sustentador principal y se escoge la más privilegiada de estas dos, agrupando en las clases sociales I y II (más privilegiadas), III, IV, V y VI (menos privilegiadas) de acuerdo con la adaptada para la Encuesta Nacional de Salud (218, 219).

Resultados

Tabla 10. Evolución en 17 años de la prevalencia de alergia según capacidad económica.

	Ambos sexos							Varones							Mujeres						
	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
2001	8,03	9,61	7,25	8,62	7,79	7,72	8,50	6,63	8,39	6,29	7,94	6,20	5,69	6,62	9,34	10,79	8,19	9,31	9,35	9,60	9,85
2003	9,83	10,96	11,45	9,49	9,76	9,48	8,96	8,72	10,67	11,16	8,10	8,46	8,71	7,24	10,90	11,21	11,74	10,76	11,14	10,18	10,45
2015	10,54	12,20	12,63	11,29	10,12	8,32	9,53	8,91	13,00	10,97	9,30	8,00	5,77	7,73	12,10	11,32	14,30	13,26	12,30	10,41	11,12
2011	10,74	12,20	13,96	11,39	9,53	9,72	11,26	9,66	11,25	14,49	10,62	7,68	8,95	9,11	11,77	13,26	13,49	12,06	11,77	10,50	12,91
2014	13,37	15,06	15,62	13,58	13,44	12,47	12,87	11,25	13,45	14,05	11,68	10,72	10,28	10,32	15,38	16,70	17,08	15,41	16,32	14,61	15,03
2017	15,28	16,69	17,60	14,72	16,47	13,77	15,80	13,23	16,32	15,09	13,72	15,54	10,82	12,36	17,22	17,06	19,95	15,69	17,57	16,58	18,61

La distribución es normal (Shapiro-Wilk $W = 0,99201$, $p\text{-value} = 0,7979$) y la aplicación del test de Student nos resulta significativo, tomando los datos del primer año de medición y el último, con una $p < 2,2 \cdot 10^{-16}$.

En la figura 15 se visualizan los datos cuantitativos de la tabla 10, es decir niveles superiores de alergia de unos dos puntos en los niveles de renta más altos con respecto a los niveles de renta más bajos. Esto es independiente del año en que se realiza la medición. La capacidad económica más alta hace que la alergia suba en general hasta dos puntos, más claramente en varones. En varones en I y II esta por encima hasta 6-7 puntos (por ejemplo en 2005). El nivel II es el más alto de alergia en todos y se acentúa sobre todo de 2014 a 2017.

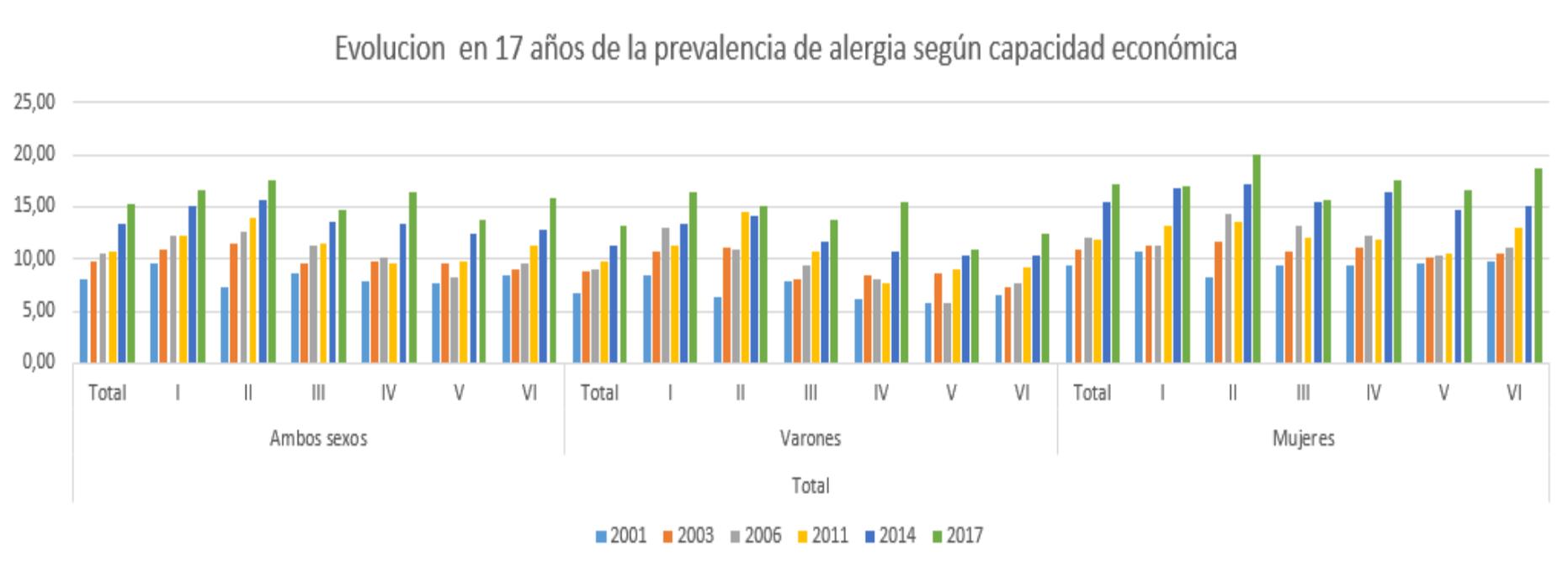


Figura 15. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abscisas los diferentes niveles económicos con 3 divisiones: cifras totales en ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Resultados

La tabla 11 nos muestra la estadística descriptiva de la evolución de la prevalencia de alergia por sexo y nivel económico. El promedio en 17 años muestra valores de dos puntos por encima en los niveles de renta más altos, tanto en cifras totales como en varones siendo en mujeres menos marcado. La mediana sigue valores similares a la media mostrando una distribución centrada hecho que corroboran tanto la curtosis como el coeficiente de asimetría. En varones en el nivel de renta más bajo la cifra es de 8,36% y en uno de los niveles más altos (el II llega a ser de 12,18%). En mujeres es menos marcado pero el punto más alto está en el nivel III con 14,12%.

Tabla 11. Estadística descriptiva de la evolución de la prevalencia de alergia en 17 años con respecto al nivel económico y sexo.

	Ambos sexos						Varones						Mujeres								
	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
Media	11,29799	12,78639	13,08556	11,5148	11,184	10,24586	11,15472	9,733578	12,1804	12,01023	10,22595	9,431423	8,367747	8,897966	12,78549	13,39014	14,12633	12,74887	13,07445	11,97846	11,97846
Error típico	1,061641	1,072378	1,466591	0,950357	1,296821	0,971358	1,140165	0,929204	1,108865	1,3454	0,914115	1,359305	0,894758	0,8822	1,20106	1,157846	1,672486	1,036295	1,299404	1,179234	1,179234
Mediana	10,6403	12,19984	13,29559	11,34142	9,938884	9,599768	10,39659	9,288104	12,12605	12,60703	9,959522	8,230522	8,827113	8,4227	11,93617	12,28735	13,89698	12,65887	12,03526	10,45214	10,45214
Desviación estándar	2,600478	2,626778	3,592399	2,327889	3,176549	2,379331	2,792823	2,276075	2,716153	3,295545	2,239114	3,329604	2,191701	2,160939	2,941985	2,836132	4,096738	2,538394	3,182877	2,888522	2,888522
Varianza de la muestra	6,762486	6,899965	12,90533	5,419067	10,09046	5,661216	7,799862	5,180519	7,377488	10,86061	5,013633	11,08626	4,803553	4,669658	8,655276	8,043645	16,78326	6,443444	10,1307	8,343561	8,343561
Curtosis	-0,310822	-0,786776	0,582826	-1,221875	0,293854	-1,147669	0,117684	0,261387	0,286637	0,964538	-0,637388	2,378517	-1,808573	-0,330454	-0,778045	-2,054681	-0,051316	-1,522796	-1,281499	-0,707789	-0,707789
Coefficiente de asimetría	0,566454	0,541411	-0,609606	0,217874	1,042163	0,687773	0,988049	0,375021	0,220538	-1,136058	0,623238	1,515955	-0,424028	0,789159	0,655596	0,643792	0,001202	-0,080332	0,589773	1,112356	1,112356
Rango	7,246902	7,076592	10,35347	6,100556	8,679832	6,05157	7,299286	6,604068	7,927357	8,80182	5,77976	9,337357	5,126769	5,740558	7,875337	6,266218	11,75172	6,3748	8,217235	6,98648	6,98648
Mínimo	8,03031	9,612856	7,248917	8,621231	7,789232	7,719569	8,499408	6,627951	8,390137	6,293031	7,937338	6,198306	5,689424	6,622362	9,342314	10,7934	8,193646	9,310764	9,35117	9,596413	9,596413
Máximo	15,27721	16,68945	17,60239	14,72179	16,46906	13,77114	15,79869	13,23202	16,31749	15,09485	13,7171	15,53566	10,81619	12,36292	17,21765	17,05961	19,94536	15,68556	17,56841	16,58289	16,58289
Suma	67,78792	76,71836	78,51337	69,0888	67,10398	61,47517	66,92833	58,40147	73,08241	72,06139	61,35572	56,58854	50,20648	53,3878	76,71296	80,34084	84,758	76,4932	78,44673	71,87075	71,87075

La figura 16 nos muestra gráficamente los resultados de la anterior tabla, siendo lo más notable el aumento en los niveles más altos, hecho apreciable visualmente, siendo de mayor cuantía en mujeres. Vemos en la figura que hay un descenso claro desde los niveles económicos más altos a los de menor nivel tanto en ambos sexos como en cada uno de ellos por separado.

Resultados

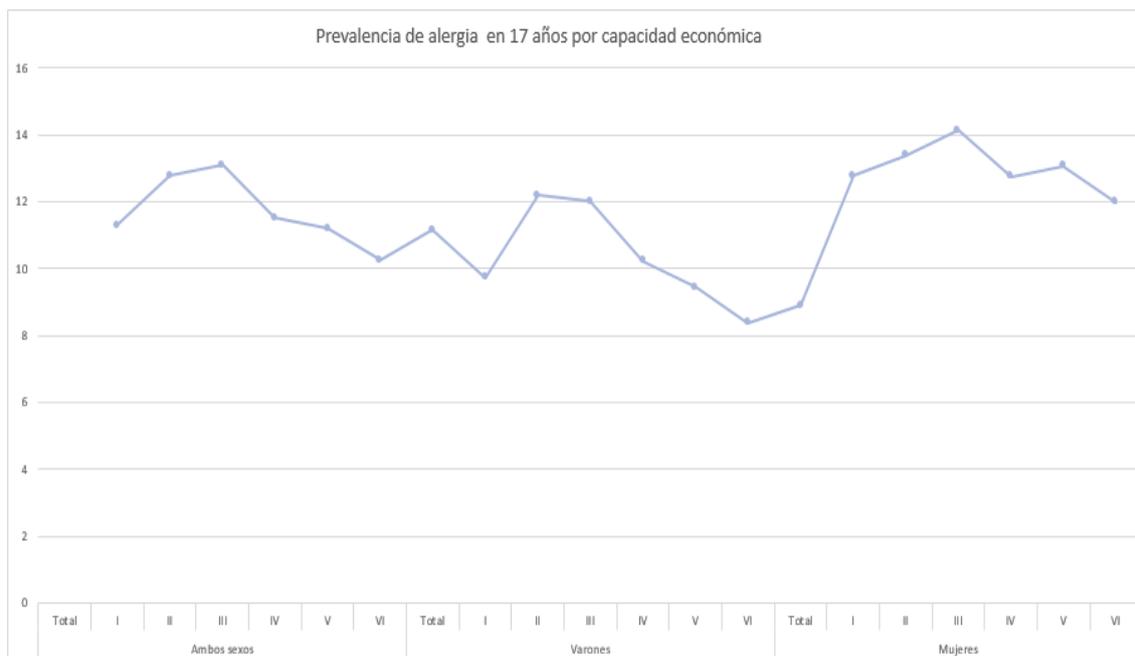


Figura 16. Promedio de evolución de prevalencia de alergia en 17 años, representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos niveles económicos diferenciándose por cifras totales y sexo.

Similares resultados observamos en la figura 17 visualizando que esto es independiente del año de medición, aunque la disminución, en los niveles económicamente más bajos, es más profunda en varones en los últimos años de medición.

Se aplicó el test de Student que mostró no significatividad en las diferencias encontradas en la prevalencia entre los niveles más altos y más bajos, $t = 1,9754$, $df = 21,35$, $p\text{-valor} = 0,06129$. El valor rozaba la significatividad pero no llegaba a dicha condición.

Previamente realizamos el estudio de las varianzas con el test de Barlett con un p valor de 0,5683. Si en el test de Barlett la $p > 0,05$ se acepta la hipótesis nula y las varianzas son iguales o sea homocedásticas.

Resultados

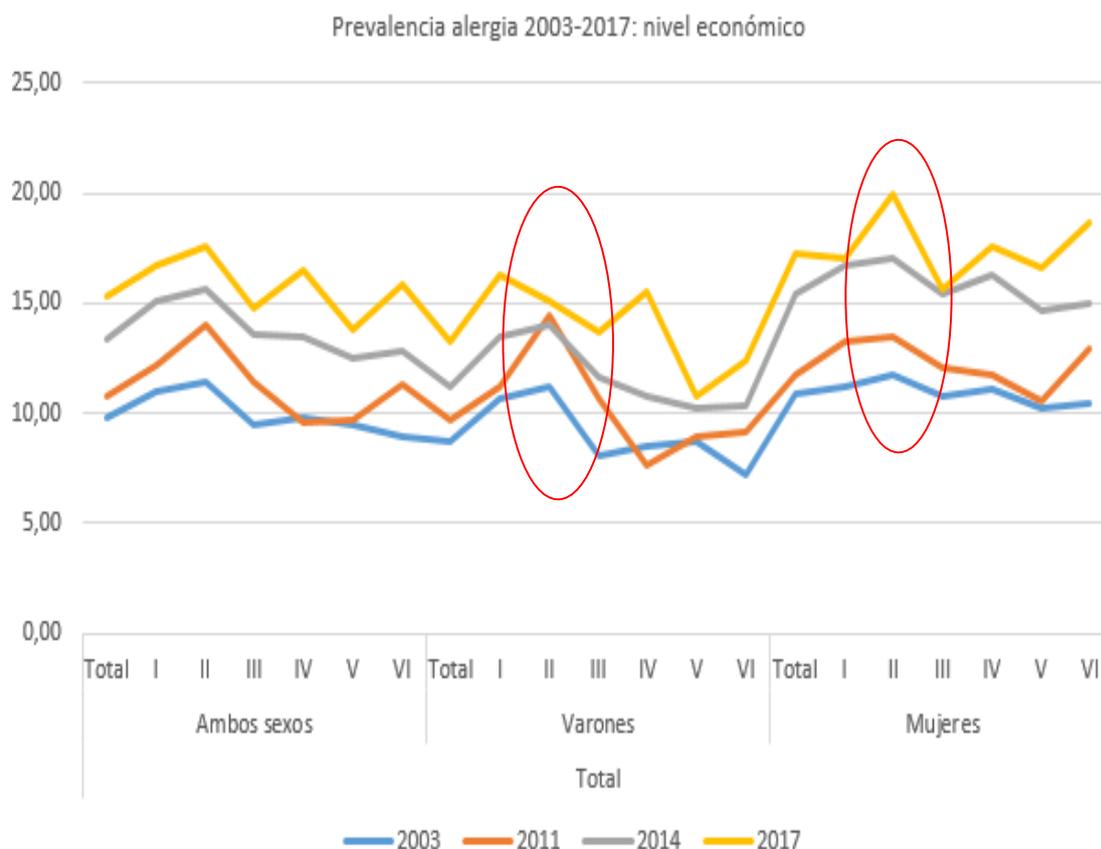


Figura 17. Serie temporal en 15 años de prevalencia de alergia representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos niveles económicos diferenciándose por cifras totales de ambos sexos y desglosados en varones y mujeres. Con desglose según el año de medición que se expresa en diferentes colores según leyenda de pie de figura.

La prevalencia en varones es más alta en las capacidades económicas más altas (sobre todo en el grupo II) mientras que en los grupos V y VI disminuye hasta 6-7 puntos. En las mujeres hay más homogeneidad entre grupos pero los puestos más altos los ocupa el grupo II.

Prevalencia alergia: capacidad económica, varones



Figura 18. Serie temporal en 15 años de prevalencia de alergia en varones representando el porcentaje de prevalencia en ordenadas y en abcisas los años de medición, diferenciando los distintos niveles económicos según figuras geométricas expresadas en leyenda del margen de la figura.

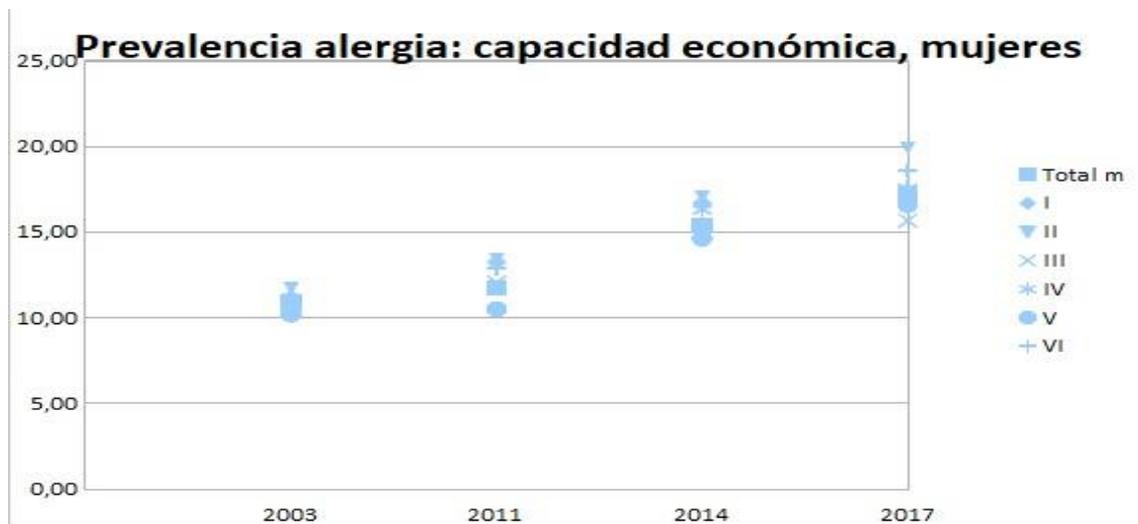


Figura 19. Serie temporal en 15 años de prevalencia de alergia en mujeres representando el porcentaje de prevalencia en ordenadas y en abcisas los años de medición, diferenciando los distintos niveles económicos según figuras geométricas expresadas en leyenda del margen de la figura.

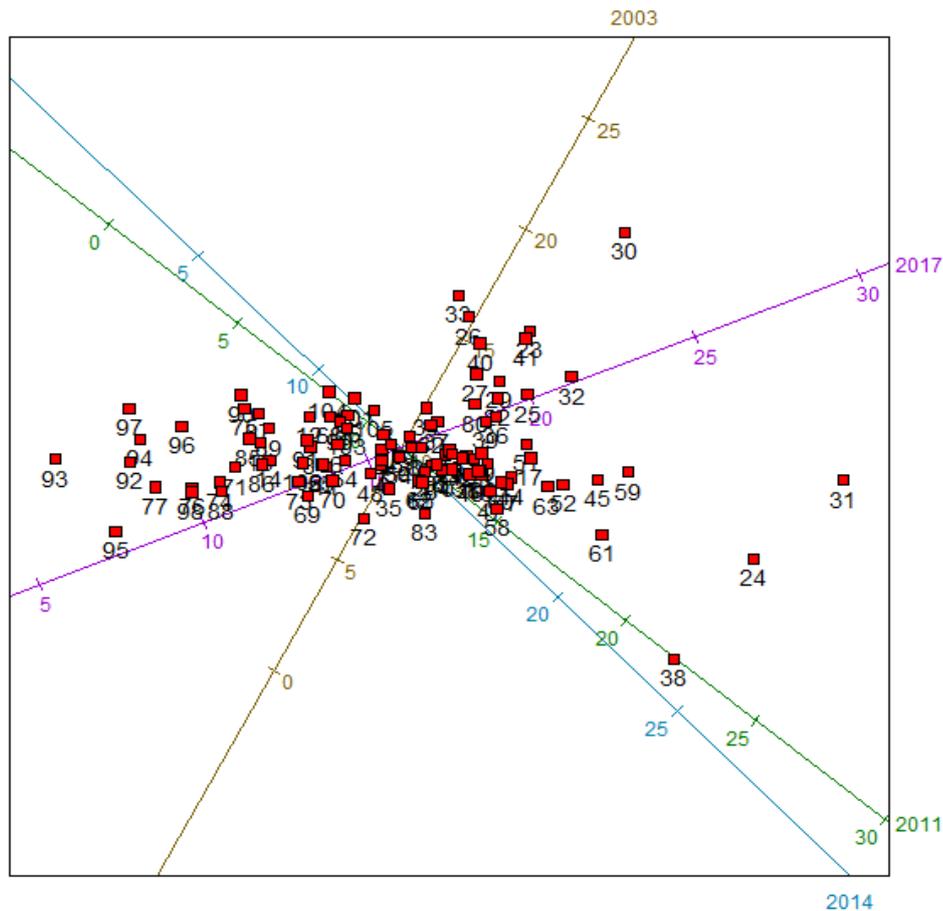


Figura 20. Biplot de representación de la distribución de la prevalencia de alergia por niveles económicos, situándose cada uno de los puntos de la muestra ordenados por edad, sexo y niveles según la tabla 12, siendo los ejes los años en que se toman las muestras.

La figura 20 nos muestra el biplot que representa la distribución de la gran cantidad de variables. La tabla 12 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 20 y que corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de variables de edad y sexo. Las variables 30-31 corresponden a varones de 15-24 años de los niveles más altos, el 38 el de mujeres del mismo nivel económico y edad junto con el 24 que es el mismo grupo de edad del mismo nivel en ambos sexos; estarían en la zona más separada proyectándose a los niveles más altos de prevalencia. La mayoría de observaciones reflejaría valores similares, y ya que la proximidad entre puntos se interpreta en términos de similitud, el abigarramiento es visualmente lo que llama la atención.

Resultados

Tabla 12. Clave numérica para identificar cada una de las variables que aparecen en la figura 20 y que corresponden a diferentes nichos de edad y sexo junto con nivel económico.

1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18		19		20	
Total																																							
Ambos sexos										Varones						Mujeres																							
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI																			
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41																				
15-24 años																																							
Ambos sexos										Varones						Mujeres																							
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI																			
43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62																				
25-44 años																																							
Ambos sexos										Varones						Mujeres																							
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI																			
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83																				
45-64 años																																							
Ambos sexos										Varones						Mujeres																							
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI																			
85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104																				
65 y más años																																							
Ambos sexos										Varones						Mujeres																							
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI																			

4.1.5. Estadística descriptiva de la prevalencia de alergia según tipo de ocupación

La evolución en 25 años de la prevalencia de alergia según ocupación no muestra grandes diferencias según el grupo: ocupado, parado, inactivo. En parados hay una ligera ventaja que disminuye en las últimas mediciones. En varones en los últimos años la tendencia a que haya más alergia en parados disminuye.

La distribución sigue un patrón que ya no se distribuye con normalidad, como en los apartados anteriores, puesto que algunos nichos de edad y sexo difieren en gran manera (Shapiro-Wilk $W = 0,92766$, $p\text{-value} = 0,00158$). Al ser la $p < 0,05$ se rechaza la hipótesis nula de normalidad de la distribución. En el test de Welch se halló no significatividad con $p\text{-valor} = 0,9517$.

Tabla 13. Evolución en 25 años de la prevalencia de alergia según ocupación laboral.

	EVOLUCION EN 25 AÑOS DE LA ALERGIA SEGÚN OCUPACION LABORAL											
	Ambos sexos				Varones				Mujeres			
	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
1993	6,25	5,28	6,58	6,92	5,32	4,68	5,03	6,48	7,14	6,68	8,56	7,13
1995	8,05	7,06	8,54	8,67	6,83	6,10	8,01	7,52	9,19	9,12	9,12	9,24
1997	7,95	6,07	10,13	8,90	6,18	4,96	8,15	7,28	9,60	8,18	12,17	9,76
2001	8,03	7,69	8,51	8,26	6,63	6,46	8,21	6,61	9,34	9,86	8,74	9,14
2011	10,72	11,17	11,08	10,10	9,62	10,34	9,19	8,72	11,77	12,18	13,52	10,98
2014	13,32	14,59	14,06	11,65	11,16	12,68	10,61	9,12	15,38	17,01	17,47	13,34
2017	15,18	15,56	14,94	14,79	13,09	13,68	12,42	12,44	17,15	17,83	17,29	16,45

Tenemos la visualización de los datos cuantitativos de la tabla 13 en la figura 21 con la serie temporal de evolución de la prevalencia de alergia según ocupación laboral en 25 años, representando en ordenadas el porcentaje de prevalencia y en abcisas cada uno de los años de medición y diferenciando los distintos grupos de ocupación según colores de la leyenda. Observamos que excepto en 2017 los porcentajes no difieren excesivamente según ocupación y lo único que asciende en la gráfica es la lógica evolución al alza del porcentaje de alergia con el paso de los años. En 1993 en cifra total de ambos sexos era de 6,25% y en total ser mujer parada representaba el 8,56%. El parado y el inactivo tenían una tendencia en ambos sexos que era de mayor incidencia de alergia que el ocupado. Pero esa tendencia se rompe desde

2011 pasando a tener más alergia los ocupados. Esta tendencia se replica tanto en hombres como en mujeres.

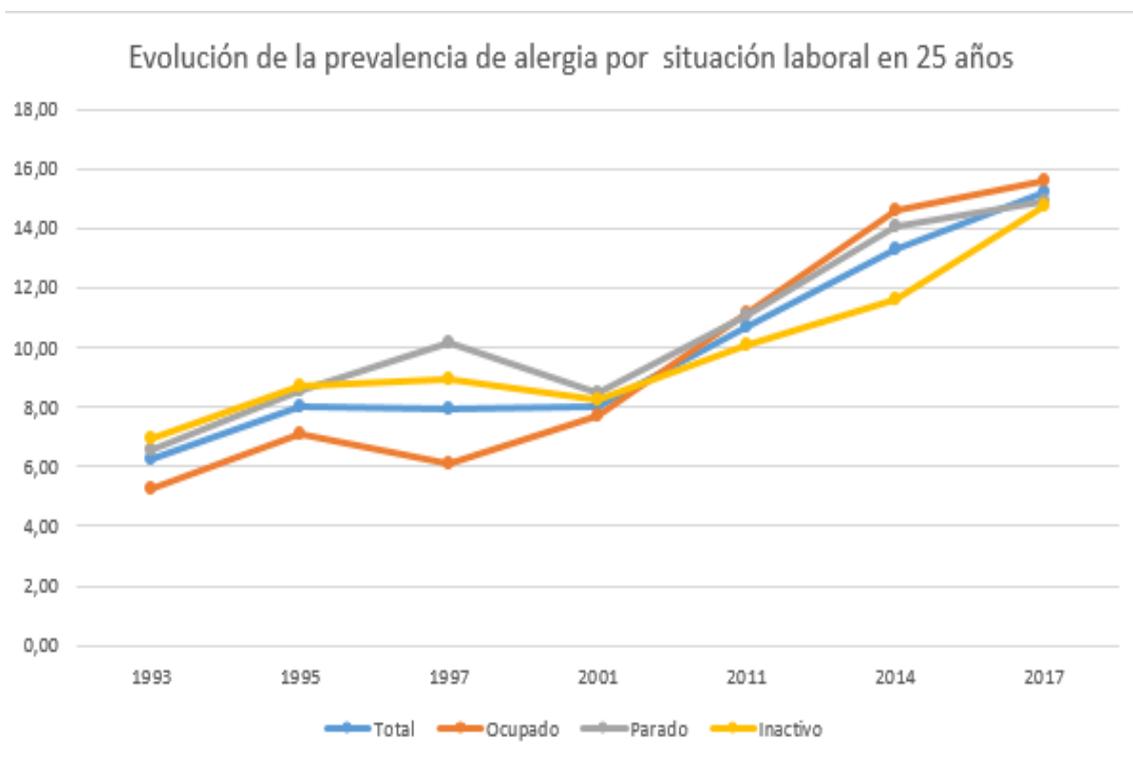


Figura 21. Serie temporal de evolución de la prevalencia de alergia según ocupación laboral en 25 años, representando en ordenadas el porcentaje de prevalencia y en abcisas cada uno de los años de medición y diferenciando los distintos grupos de ocupación según colores de la leyenda.

En la figura 22 tenemos la visualización en forma de diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

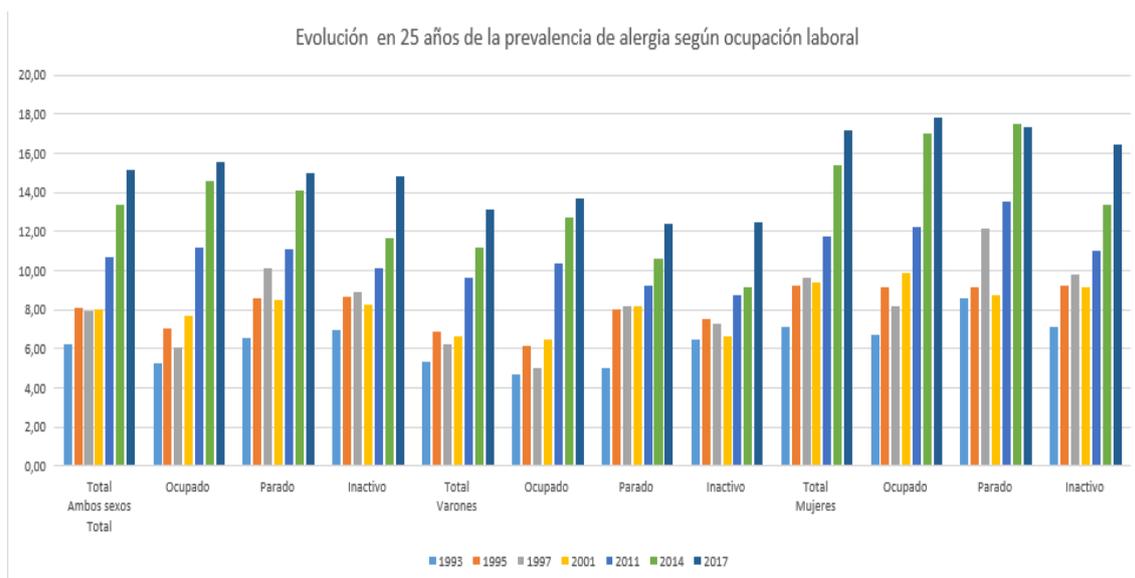


Figura 22. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

La tabla 14 nos muestra la estadística descriptiva de la evolución de la prevalencia de alergia por tipo de ocupación y sexo. En cifras totales hay una ligera ventaja en el grupo de los parados. Hay unas ligeras diferencias de más en parados sobre todo en mujeres. La población inactiva y los ocupados muestran cifras muy parecidas, tanto en varones como mujeres.

La media muestra valores más altos, tanto en cifras totales como en mujeres siendo en varones menos marcado. La mediana sigue valores similares a la media mostrando una distribución centrada hecho que corroboran tanto la curtosis como el coeficiente de asimetría. Las diferencias entre la desviación oscilan entre un 20 y un 40%.

Resultados

Tabla 14. Estadística descriptiva de la evolución de la prevalencia de alergia en 25 años con respecto a situación laboral y sexo.

	EVOLUCION EN 25 AÑOS DE LA ALERGIA SEGÚN OCUPACION LABORAL											
	Ambos sexos				Varones				Mujeres			
	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
Media	9,92746	9,632612	10,54665	9,895845	8,404855	8,41517	8,800824	8,313302	11,36679	11,55165	12,4114	10,86378
Error típico	1,237321	1,574449	1,154748	0,989053	1,103259	1,419877	0,875006	0,783368	1,376034	1,644112	1,462725	1,176551
Mediana	8,051865	7,690292	10,12839	8,896573	6,828781	6,455338	8,205574	7,522936	9,601729	9,859276	12,17391	9,763171
Desviación estándar	3,273643	4,1656	3,055176	2,616788	2,918949	3,756642	2,315049	2,072598	3,640644	4,349912	3,870006	3,11286
Varianza de la muestra	10,71674	17,35222	9,334098	6,847581	8,520266	14,11236	5,35945	4,295661	13,25429	18,92174	14,97694	9,689898
Curtosis	-0,862831	-1,645825	-1,085988	1,273308	-1,071555	-1,900192	0,866365	2,482475	-0,76423	-1,254699	-1,771552	0,688241
Coefficiente de asimetría	0,752411	0,58729	0,400242	1,154953	0,712462	0,518069	-0,026113	1,528078	0,76592	0,673532	0,422936	0,982691
Rango	8,930033	10,27708	8,360193	7,868035	7,769586	9,002977	7,392648	5,956173	10,00766	11,15394	8,909661	9,32269
Mínimo	6,245557	5,282983	6,577956	6,918199	5,321818	4,680011	5,026939	6,484497	7,1411	6,68106	8,563257	7,13005
Máximo	15,17559	15,56006	14,93815	14,78623	13,0914	13,68299	12,41959	12,44067	17,14876	17,835	17,47292	16,45274
Suma	69,49222	67,42828	73,82654	69,27092	58,83398	58,90619	61,60577	58,19311	79,56755	80,86153	86,8798	76,04645



Figura 23. Promedio de evolución de prevalencia de alergia en 25 años, representando en ordenadas el porcentaje de prevalencia y en abcisas los distintos tipos de ocupación diferenciándose por cifras totales y sexo.

Teniendo en cuenta la evolución en 25 años las cifras medias describen un moderado aumento en parados tanto en hombres como mujeres.

Resultados

Tabla 15. Clave numérica con todos los nichos de edad, sexo y ocupación.

1												2												3												4												5												6												7												8												9												10												11												12											
Ambos sexos												Total												Varones												Mujeres																																																																																																											
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo																																																																																																				
13												14												15												16												17												18												19												20												21												22												23												24											
Ambos sexos												Total												Varones												Mujeres																																																																																																											
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo																																																																																																
25												26												27												28												29												30												31												32												33												34												35												36											
Ambos sexos												Total												Varones												Mujeres																																																																																																											
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo																																																																																																
37												38												39												40												41												42												43												44												45												46												47												48											
Ambos sexos												Total												Varones												Mujeres																																																																																																											
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo																																																																																																
49												50												51												52												53												54												55												56												57												58												59												60											
Ambos sexos												Total												Varones												Mujeres																																																																																																											
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo																																																																																																

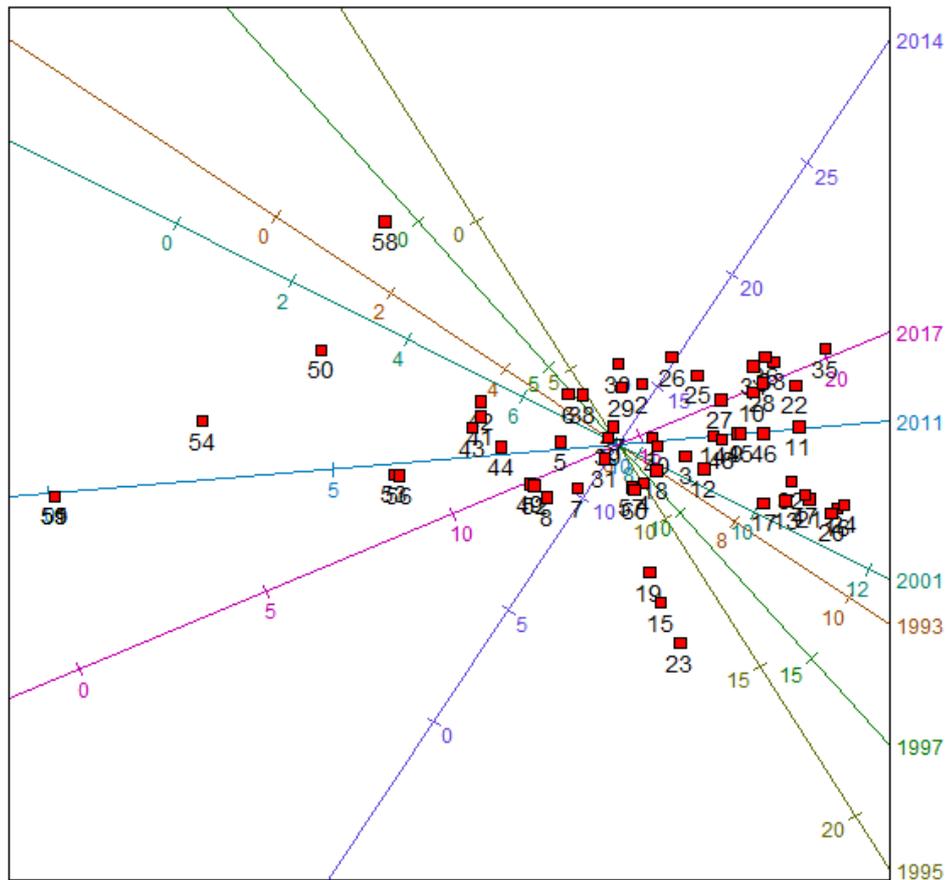


Figura 24. Biplot representado la evolución de la prevalencia de alergia en los años indicados en ejes según tipo de ocupación.

El biplot de la figura 24 representa la evolución de la prevalencia de alergia en los años indicados en ejes según tipo de ocupación, y ocupando las posiciones más extremas se observan los nichos 50, 54 y 58 que corresponde a grupo de 65 y más años.

La tabla 15 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 24 y que corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de variables de edad y sexo. La proximidad entre puntos se interpreta en términos de similitud, el ángulo en términos de correlación. Nula correlación entre 2014 y el resto de años con vectores que son perpendiculares, mientras que se presentan superpuestos 2011-2017 y 1993-2001 con variables más relacionadas y ejes con proximidad a 180° con lo que hay una correlación positiva alta.

4.1.6. Estadística descriptiva de la prevalencia de alergia según tipo de jornada laboral

Según el tipo de jornada laboral obtenemos resultados dispares en cuanto a la alergia porque en algunos casos, en determinadas jornadas, el número de personas es más bajo y los porcentajes de prevalencia menos significativos.

El test de Welch no es significativo con p-valor = 0,06144, aunque roza el límite con la significatividad no le podemos atribuir dicha condición. La distribución sigue un patrón que ya no se distribuye con normalidad puesto que algunos nichos de edad y sexo difieren en gran manera con el resto (Shapiro-Wilk $W = 0,91514$, p-valor = 0,00013, se rechaza la hipótesis nula y no hay distribución normal).

Resultados

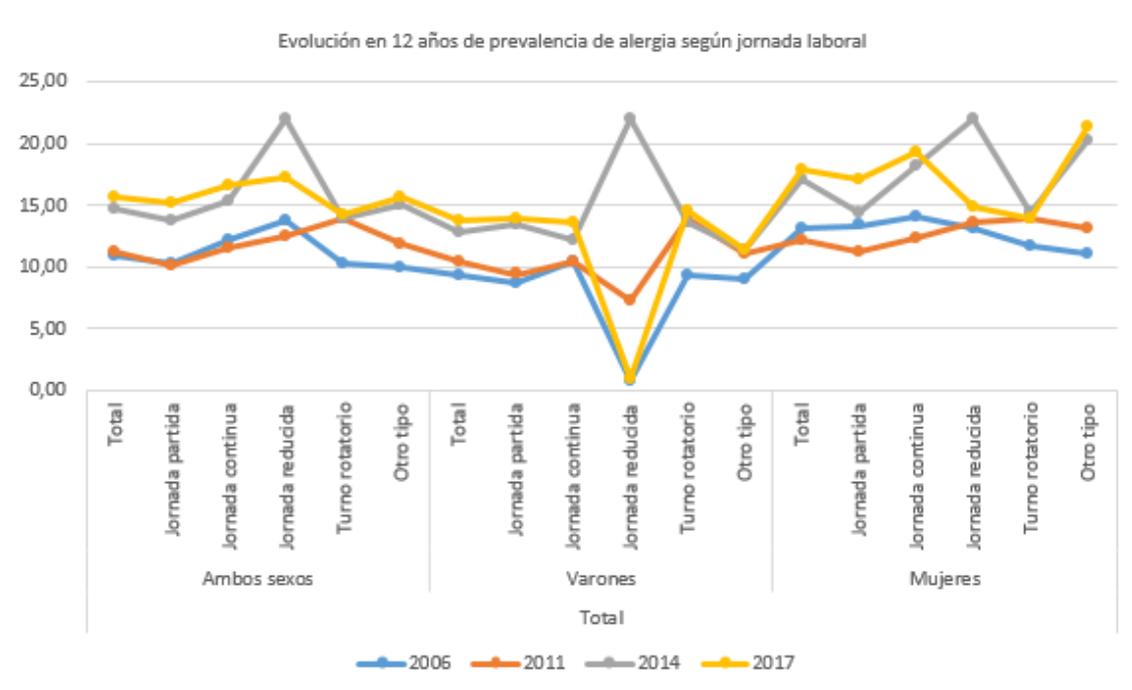


Figura 25. Evolución de la prevalencia de alergia según jornada laboral, representando en abscisas cada uno de los tipos de jornada subdivida por sexos y con distintos colores el año de medición. En ordenadas se muestra la cifra de prevalencia.

Tabla 16. Prevalencia de alergia en 12 años según jornada laboral.

	Total																	
	Ambos sexos						Varones						Mujeres					
	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
2006	10,86	10,21	12,23	13,75	10,28	9,89	9,35	8,74	10,42	0,82	9,38	9,07	13,07	13,36	14,02	13,14	11,73	11,02
2011	11,22	10,12	11,48	12,46	13,96	11,86	10,39	9,40	10,50	7,33	13,95	10,99	12,23	11,30	12,33	13,55	13,98	13,05
2014	14,64	13,79	15,41	21,99	13,91	15,06	12,76	13,43	12,25	22,03	13,64	11,30	17,03	14,41	18,25	21,97	14,32	20,18
2017	15,61	15,18	16,55	17,20	14,24	15,64	13,73	13,99	13,58	0,91	14,51	11,45	17,90	17,16	19,32	14,90	13,88	21,38

Resultados

Tabla 17. Estadística descriptiva de la evolución de la prevalencia de alergia en 12 años con respecto a tipo de jornada laboral.

	2006	2011	2014	2017
Media	10,76638	9,852179	12,45421	12,58155
Error típico	0,555165	0,524488	0,792118	0,78563
Mediana	11,10879	11,41018	14,28682	14,89013
Desviación	4,710735	4,450427	6,721347	6,666295
Varianza de	22,19103	19,8063	45,17651	44,43948
Curtosis	1,340405	0,823071	-0,163221	-0,123681
Coefficiente	-0,337182	-1,357503	-0,795558	-1,108564
Rango	21,84294	15,84757	24,62017	22,37965

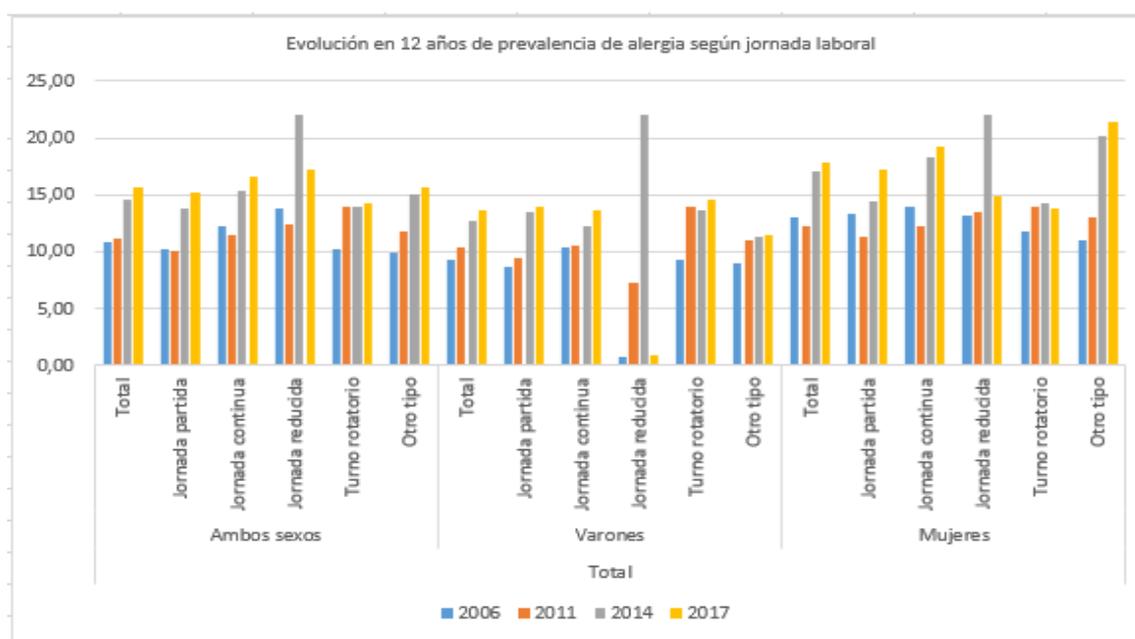


Figura 26. Diagrama de barras con el porcentaje de prevalencia de alergia en ordenadas y en abcisas los diferentes tipos de jornada con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Resultados

Tabla 18. Clave de alergia y tipo de jornada laboral con todos los nichos de edad, sexo y ocupación.

Total																	
Ambos sexos						Varones						Mujeres					
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
16-24 años																	
Ambos sexos						Varones						Mujeres					
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
25-44 años																	
Ambos sexos						Varones						Mujeres					
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
45-64 años																	
Ambos sexos						Varones						Mujeres					
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo

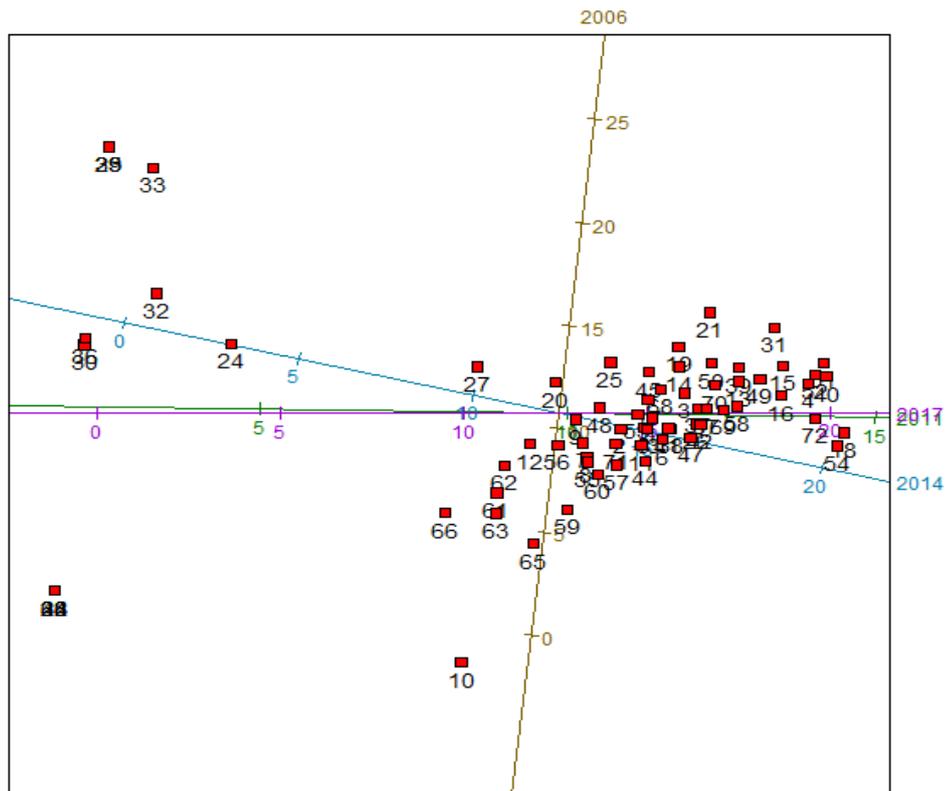


Figura 27. Biplot de representación de prevalencia de alergia por tipo de jornada.

La figura 27 es un biplot representando la evolución de la prevalencia de alergia en los años indicados en ejes según tipo de jornada laboral, y ocupando las posiciones más extremas y aisladas 28, 32 y 33 que corresponden al grupo de 15-24 años. La tabla 18 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 27 y que corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de variables de edad y sexo. La proximidad entre puntos se interpreta en términos de similitud, de ahí la gran concentración en una de las zonas de la gráfica. El acercamiento entre los ejes, ya que el ángulo nos comunica correlación, sería sobre todo en los resultados de los últimos años quedando menos homogénea la primera medición que es la de 2006.

4.1.7. Estadística descriptiva de la prevalencia de asma

Para realizar el análisis estadístico de la prevalencia de asma, hemos escogido 4 fechas de medición: 2006, 2011, 2014, 2017. Esto es debido a que, es a partir de 2006 cuando el INE separa los datos de asma de los datos de bronquitis crónica y enfisema, los cuales eran recogidos conjuntamente hasta ese momento. La prevalencia de asma no ha sufrido la misma evolución que la prevalencia de alergia. Como vemos en los sucesivos desarrollos, las cifras de asma se han mantenido estables en los últimos 12 años.

Teniendo en cuenta la cifra total en ambos sexos, el asma de 2006 a 2017 pasa de una prevalencia de 4,45% a 4,68%.

En la población de más de 75 años es de 7,77% en 2006 y de 6,51% en 2017. En dicha población en varones llega a ser de 8,94% en 2006. En varones hay aumentos de hasta 6 puntos en 2006 en más de 75 años con respecto el resto de nichos. En mujeres menos acusado pero también hay un aumento. En las poblaciones de más de 75 años se alcanzan cifras de 6% que contrasta con las cifras en varones en las poblaciones de 35-44 y de 55-64 años que son 2 puntos menores. En mujeres las cifras son más altas en todos los segmentos de edad. La conclusión es el evidente aumento de la prevalencia en poblaciones de más edad. El aumento en la población de más de 75 años tanto en cifras totales como particularmente en varones y mujeres es evidente, así como la caída en grupos de edades medias (45-54 años y

55-64 años). Llama la atención no solo la mayor prevalencia en las poblaciones de más edad, sino también la uniformidad en los 4 puntos de medición a lo largo de los 12 años.

La T de Student entre el primer año de medición y el último no es significativa con p-valor= 0,894. La distribución cumple criterios de normalidad (Shapiro-Wilk W = 0,93695, p-valor = 0,1395, no se rechaza la hipótesis nula de normalidad).

Tabla 19. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (15-24 años, 25-34 años, 35-44 años, 45-54 años, 55-64 años, 65-74 años, más de 75 años).

		2006	2011	2014	2017
Ambos sexos	Total	4,45	4,08	4,37	4,68
	15-24	5,03	4,72	3,95	5,22
	25-34	4,47	4,49	4,80	5,31
	35-44	2,97	4,03	4,71	4,19
	45-54	3,35	2,97	3,02	3,98
	55-64	4,06	3,38	4,19	3,78
	65-74	5,68	4,33	5,06	4,79
	75 y más años	7,77	5,28	5,42	6,51
Varones	Total	3,93	3,38	3,90	3,74
	15-24	4,72	4,67	3,81	3,45
	25-34	3,79	3,61	5,11	4,31
	35-44	2,29	4,12	3,91	4,36
	45-54	2,67	1,80	2,46	2,70
	55-64	3,61	2,28	2,93	2,65
	65-74	5,14	3,18	4,15	4,15
	75 y más años	8,94	4,45	6,22	5,38
Mujeres	Total	4,94	4,75	4,82	5,57
	15-24	5,35	4,76	4,10	7,08
	25-34	5,20	5,40	4,49	6,30
	35-44	3,68	3,95	5,54	4,02
	45-54	4,03	4,12	3,58	5,25
	55-64	4,48	4,42	5,40	4,86
	65-74	6,09	5,32	5,86	5,36
	75 y más años	6,94	5,82	4,89	7,25

Resultados

Tabla 20. Valores de estadística descriptiva a lo largo 12 años (2006, 2011, 2014, 2017) sobre la prevalencia de asma, distribuidos por divisiones por totales de ambos sexos, varones y mujeres desglosados en grupos de edad (15-24 años, 25-34 años, 35-44 años, 45-54 años, 55-64 años, más de 75 años).

	Ambos sexos								Varones								Mujeres							
	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años
Media	4,3925934	4,7299883	4,7663263	3,9776538	3,3300291	3,8534	4,9620704	6,2442956	3,7362623	4,1625007	4,2045289	3,6693075	2,4102794	2,8688171	4,1567971	6,2453977	5,0182625	5,3238151	5,3463121	4,2973922	4,2465246	4,7894105	5,6564265	6,2263908
Error típico	0,123546	0,2801391	0,1956697	0,3649743	0,2314268	0,1791979	0,2825997	0,5781444	0,1271715	0,317857	0,3354445	0,4683276	0,2091467	0,2809448	0,3991907	0,9673038	0,1871555	0,6392484	0,3735408	0,4208412	0,355999	0,2266148	0,1909591	0,5405822
Mediana	4,408059	4,8737817	4,6437099	4,1141653	3,1878081	3,9187726	4,9222747	5,963503	3,8189376	4,2407516	4,0517427	4,0111445	2,5673124	2,7915571	4,1512949	5,7989987	4,8800008	5,0581925	5,3002643	3,9859545	4,0773431	4,66989	5,6082623	6,3804994
Desviación estándar	0,247092	0,5602783	0,3913395	0,7299486	0,4628536	0,3583958	0,5651994	1,1562888	0,2543431	0,6357141	0,670889	0,9366551	0,4182933	0,5618895	0,7983814	1,9346076	0,374311	1,2784967	0,7470816	0,8416825	0,711998	0,4532295	0,3819183	1,0811643
Varianza de la muestra	0,0610545	0,3139117	0,1531466	0,5328249	0,2142334	0,1284476	0,3194504	1,3370038	0,0646904	0,4041324	0,4500921	0,8773229	0,1749693	0,3157199	0,6374128	3,7427066	0,1401087	1,6345539	0,5581309	0,7084294	0,5069412	0,205417	0,1458616	1,1689163
Curtosis	1,1068369	1,3170898	0,7368661	1,9219986	0,9901353	-0,3842268	0,4789568	-0,8614366	1,576055	-4,440207	-0,0544416	3,2535518	2,5957778	0,6639917	1,5008779	1,5982312	3,0364091	1,3263626	1,1999658	3,4571099	2,4208341	-0,1869302	-4,1253367	-2,3785833
Coefficiente asimetría	-0,3615367	-1,2464178	1,2515409	-1,0516779	1,2929834	-0,8329076	0,3866363	0,9161498	-1,4145015	-0,2650524	0,9996947	-1,7648749	-1,6486178	0,733389	0,0413472	1,1984168	1,7303257	1,0866522	0,3585235	1,8091404	1,3100242	1,0478453	0,3091544	-0,5254183
Rango	0,5987806	1,2733006	0,8397403	1,7384188	1,0062491	0,8133347	1,3524428	2,4935237	0,5545117	1,2739752	1,4982604	2,0702336	0,9002537	1,3286328	1,9555627	4,488908	0,8197726	2,9826214	1,8136023	1,8625755	1,6744471	0,9834745	0,7791603	2,3592661
Mínimo	4,0777374	3,9495447	4,4690726	2,9719328	2,9691256	3,3813601	4,3256446	5,2783264	3,3763311	3,4472622	3,6081849	2,2923536	1,8031195	2,2817606	3,1845179	4,4473427	4,7466378	4,098127	4,4855586	3,6775421	3,5784825	4,4171939	5,3150106	4,8926491
Máximo	4,676518	5,2228453	5,3088129	4,7103516	3,9753747	4,1946948	5,6780875	7,7718501	3,9308428	4,7212374	5,1064453	4,3625871	2,7033732	3,6103934	5,1400806	8,9362507	5,5664105	7,0807483	6,299161	5,5401176	5,2529296	5,4006684	6,0941709	7,2519152
Suma	17,570373	18,919953	19,065305	15,910615	13,320117	15,4136	19,848281	24,977182	14,945049	16,650003	16,818116	14,67723	9,6411175	11,475268	16,627188	24,981591	20,07305	21,29526	21,385248	17,189569	16,986098	19,157642	22,625706	24,905563

La tabla 20 nos muestra los valores de estadística descriptiva con respecto al asma. El promedio del grupo de más de 75 años llega a ser hasta casi el doble de alguno del resto de los grupos. Las medianas se mantienen en valores similares a la media lo que habla de un tipo de distribución centrada. Las desviaciones oscilan entre un 5 y un 20% del valor de la media. La curtosis y el coeficiente de asimetría también oscilan en valores moderadamente positivos o negativos siguiendo la misma distribución en “V” que la media.



Figura 28. Promedio de la prevalencia del asma en 12 años por grupos de edad incluidos en divisiones de cifras totales, varones y mujeres.

Llama la atención el dibujo que se produce con el aumento en las poblaciones de mayor edad, así como la caída en grupos de edades medias, porque siempre se ha dicho que el asma es una enfermedad crónica por lo que se debería mantener en una subida progresiva en el tiempo y no con baches. En realidad lo que visualizamos es

una figura en “V” o en forma de valle. Lo observamos claramente en las figuras 28, 29 y 30.

La forma de la curva descrita se mantiene en las series temporales sin cambios según el año estudiado. En todos los años estudiados y en los dos sexos hay un incremento en los grupos más jóvenes con respecto a edades medias de 45-54 años, para verse superado el resultado claramente por encima de los 65 años. Es decir, no hay una subida constante y prolongada desde las primeras edades de la vida.

EVOLUCIÓN DE LA PREVALENCIA DE ASMA EN 12 AÑOS

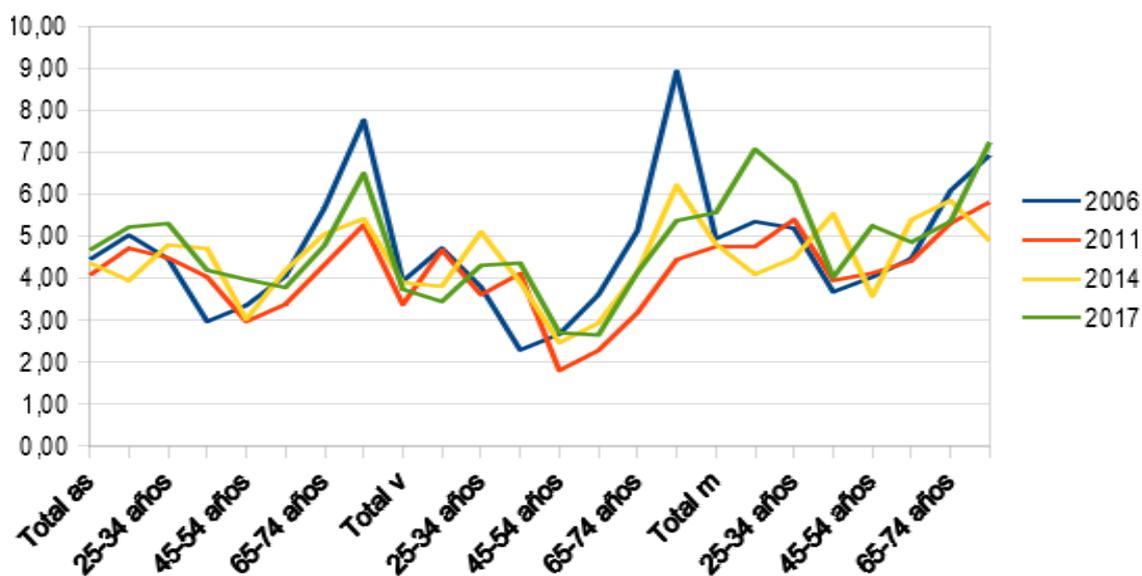


Figura 29. Desglose por años de realización de encuesta (desglosados según color de la leyenda en 4 puntos: 2006, 2011, 2014, 2017) de la evolución de la prevalencia de asma, con el porcentaje de prevalencia representado en ordenadas y en abscisas cada uno de los grupos de edad y sexo (as – ambos sexos-, v -varones- y m –mujeres-).

Evolución del asma en 12 años por grupos de edad y sexo

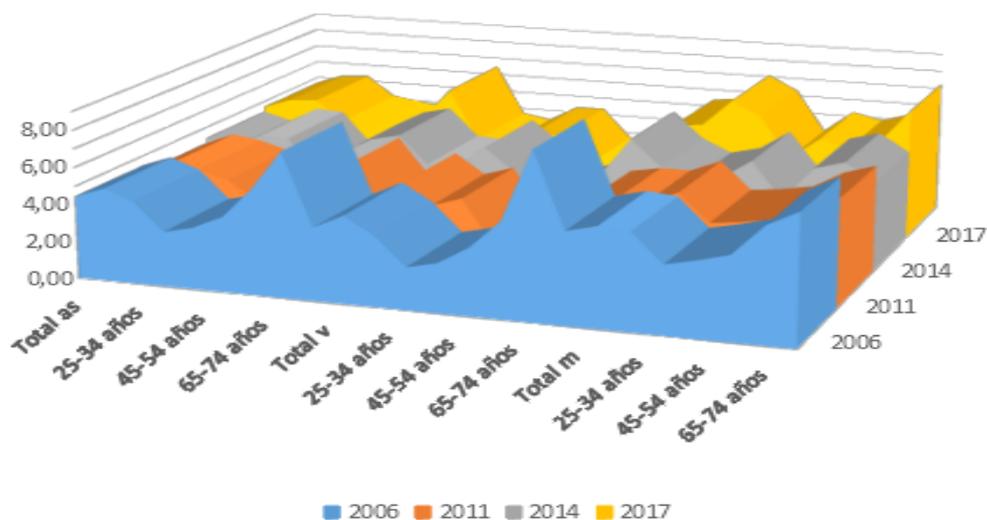


Figura 30. Dibujo en forma de relieve de la prevalencia de asma. Porcentaje de prevalencia representado en ordenadas y en abcisas cada uno de los grupos de edad y sexo (as -ambos sexos -, v -varones- y m -mujeres -). La gráfica dibuja una especie de "V" entre los grupos más jóvenes y los grupos de mayor edad.

A continuación, en la figura 31 se pone de manifiesto otra de las características del asma, el predominio femenino en todos los grupos de edad alcanzando alturas mayores las barras de prevalencia. A destacar el dato atípico de 8,94% de prevalencia en 2006 en varones del grupo de más de 75 años.

Resultados

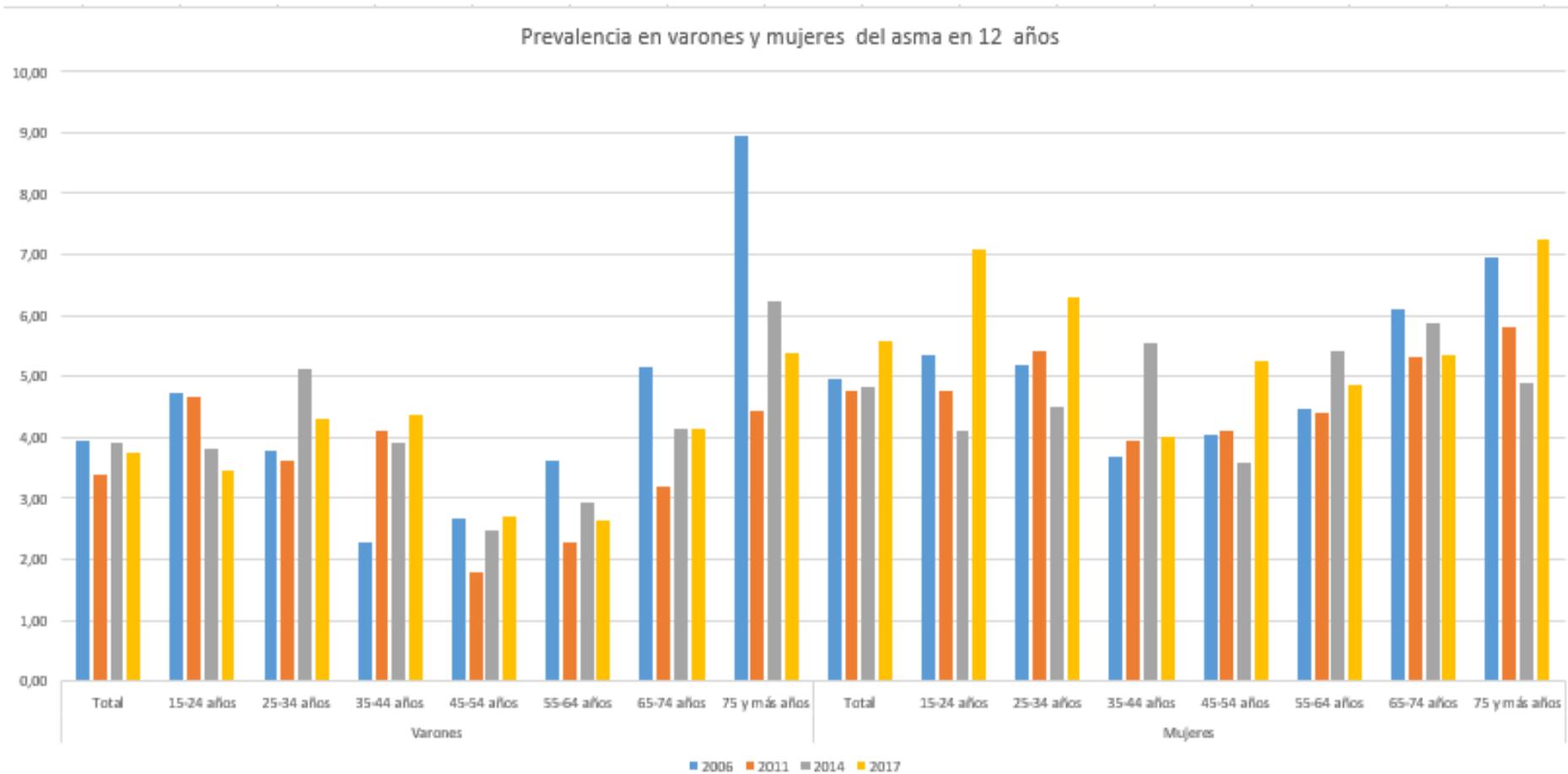


Figura 31. Diagrama de barras con el porcentaje de prevalencia de asma en ordenadas y en abcisas los diferentes grupos de edad con dos divisiones: varones y mujeres; y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Resultados

Tabla 21. Clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.

	1	2	3	4	5	6	7	8
	Ambos sexos							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	9	10	11	12	13	14	15	16
	Varones							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	17	18	19	20	21	22	23	24
	Mujeres							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	

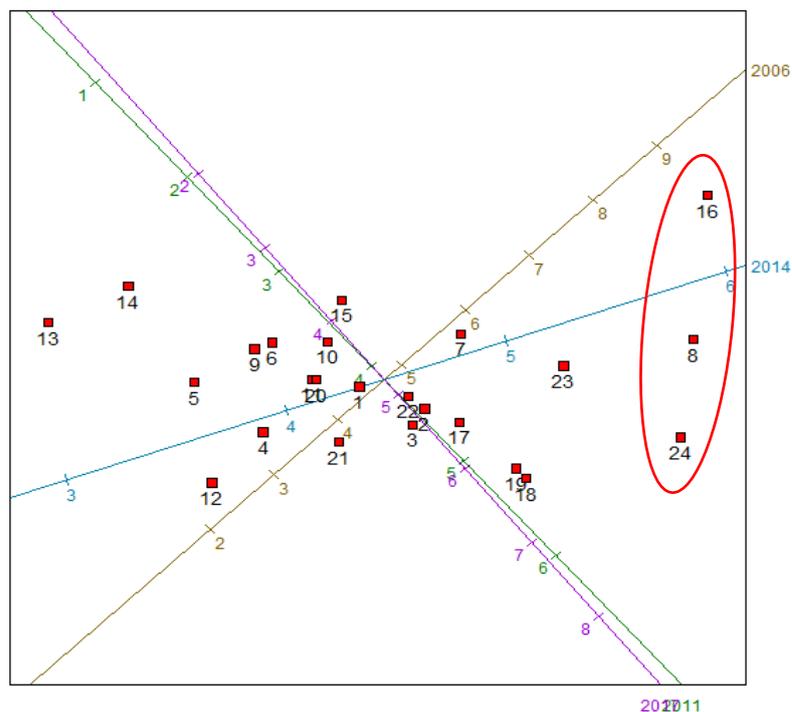


Figura 32. Biplot de la evolución de la prevalencia del asma. Representa sobre los ejes a las variables que son cada uno de los grupos de edad y sexo y los ejes son el tiempo en que se efectúan las observaciones con la divisiones representado las cifras de prevalencia.

La figura 32 representa la evolución de la prevalencia de asma en los años indicados en ejes. Se trata de representar más de 2 variables de forma que se puedan situar las observaciones para su visualización. La tabla 21 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 32 y que corresponden a diferentes nichos de edad y sexo. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de las variables de edad y sexo.

Las proyecciones de los años sobre las variables permiten obtener los valores de la muestra de datos. Los puntos 8, 16 y 24, que representa a las poblaciones de más edad, están situados en la periferia y su proyección ortogonal en los ejes se coloca en cifras altas de prevalencia.

Correlación nula entre 2006 y 2011, sirviéndonos del ángulo que forman los ejes para discernir este aspecto.

La proximidad entre puntos, como sería el caso de los nichos 4, 5, 12, 13, 14, en la zona periférica se interpreta en términos de similitud, puesto que todos se proyectan en los valores más bajos de prevalencia sobre los ejes.

4.1.8. Estadística descriptiva de la prevalencia de asma según tipo de población

Si tenemos en cuenta no solo la edad y el sexo sino también el tipo de población en cuanto al número de habitantes (menos de 10.000 habitantes, entre 10.000 y 50.000 habitantes, entre 50.000 y 100.000 habitantes, entre 100.000 y 400.000 habitantes, más de 400.000 habitantes), podemos observar la evolución de la prevalencia de asma como muestra la tablas 22 y 23.

En general, en poblaciones de más de 400.000 habitantes se dan las cifras más bajas en totales de ambos sexos y también en mujeres, aunque en 2006 en la población de 15-24 años de ambos sexos en poblaciones de más de 400.000 habitantes es de 6,11% y en las de 100.000-400.000 habitantes es de 5,82%, que contrasta con el total de ambos sexos de 4,98% y 3,76%. Las cifras más altas son en poblaciones de 100.000 a 400.000 habitantes aunque en el caso de varones también se alcanzan valores en poblaciones de hasta 10.000 habitantes.

En el grupo de 15-24 años, destaca las cifras bajas en poblaciones de 50.000 a 100.000 habitantes: 2,82%, 2,58%, 2,78% en 2011, 2014 y 2017 respectivamente.

En varones de 45-54 años destacan cifras de 1,20%, 0,23% y 1,22% en 2006, 2011 y 2014. En mujeres del mismo grupo y población (50.000 a 100.000 habitantes) cifras de 6,43%, 3,94%, 5,30% y 6,56% en 2006, 2011, 2014 y 2017 respectivamente. En mujeres del grupo de 15-24 años, en poblaciones de más de 400.000 habitantes cifras de 12,57% en 2017.

También destacan cifras muy bajas de 1,56% en varones de 25-44 años en poblaciones de más de 400.000 habitantes y sin embargo en mujeres del mismo grupo es de 3,72%, pero en mujeres de la misma edad en poblaciones de 100.000 a 400.000 habitantes es de 6,29%.

Del mismo modo, importantes valores en franjas de más de 65 años de ambos sexos en cifras totales en los años registrados. En total en ambos sexos en 2006 y con más de 65 años es 6,66%, en poblaciones de 10.000-50.000 es de 8,10%, y en las de 100.000-400.000 de 7,09%.

A diferencia de la alergia, los máximos están en los grupos de más edad tanto en hombres como mujeres. En mujeres de 15-24 años destaca el punto máximo que corresponde a 2017 en poblaciones de 100.000-400.000 habitantes que es de 8,12% y de más de 400.000 habitantes que es de 12,57%. También vemos que en 2006 hay una clara elevación en gente de más de 65 años independiente del tipo de población pero que ha ido bajando con los años.

Resultados

Tabla 22. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (cifras totales, 15-24 años, 25-44 años).

		2006	2011	2014	2017			2006	2011	2014	2017			2006	2011	2014	2017				
Total	Ambos sexos	Total	4,45	4,08	4,37	4,68	Ambos sexos	Total	5,03	4,72	3,95	5,22	Ambos sexos	Total	3,74	4,25	4,75	4,66			
		Hasta 10.000 habitantes	4,97	4,17	4,24	4,58		Hasta 10.000 habitantes	6,62	5,06	4,36	4,43		Hasta 10.000 habitantes	3,99	4,36	4,30	4,63			
		De 10.001 a 50.000 habitantes	4,06	4,56	4,42	4,81		De 10.001 a 50.000 habitantes	3,24	4,90	3,29	5,52		De 10.001 a 50.000 habitantes	2,91	4,63	4,44	5,17			
		De 50.001 a 100.000 habitantes	4,33	3,45	4,13	4,81		De 50.001 a 100.000 habitantes	3,65	2,82	2,58	2,78		De 50.001 a 100.000 habitantes	4,47	4,53	4,08	4,58			
		De 100.001 a 400.000 habitantes	4,98	4,50	4,62	4,64		De 100.001 a 400.000 habitantes	5,82	5,85	5,33	5,66		De 100.001 a 400.000 habitantes	4,97	4,35	5,07	4,69			
		Más de 400.000 habitantes	3,76	3,18	4,31	4,54		Más de 400.000 habitantes	6,11	3,99	3,93	7,07		Más de 400.000 habitantes	2,61	3,18	5,79	3,96			
		Varones	Total	3,93	3,38	3,90		3,74	Varones	Total	4,72	4,67		3,81	3,45	Varones	Total	3,07	3,88	4,42	4,34
			Hasta 10.000 habitantes	4,52	3,83	3,84		4,16		Hasta 10.000 habitantes	5,99	7,24		4,23	3,36		Hasta 10.000 habitantes	3,28	4,23	4,13	5,11
			De 10.001 a 50.000 habitantes	3,84	4,06	4,20		3,86		De 10.001 a 50.000 habitantes	3,74	5,30		3,22	4,57		De 10.001 a 50.000 habitantes	2,52	4,17	4,80	4,16
			De 50.001 a 100.000 habitantes	3,67	2,34	3,69		3,82		De 50.001 a 100.000 habitantes	4,16	1,94		3,95	3,15		De 50.001 a 100.000 habitantes	4,75	3,95	3,96	3,69
			De 100.001 a 400.000 habitantes	4,00	3,50	4,14		3,23		De 100.001 a 400.000 habitantes	5,15	4,60		5,68	3,09		De 100.001 a 400.000 habitantes	3,79	4,42	4,47	3,74
			Más de 400.000 habitantes	3,46	2,38	3,36		3,59		Más de 400.000 habitantes	4,89	2,99		1,85	2,63		Más de 400.000 habitantes	1,56	2,33	4,44	4,38
	Mujeres	Total	4,94	4,75	4,82	5,57	Mujeres	Total	5,35	4,76	4,10	7,08	Mujeres	Total	4,45	4,63	5,08	4,98			
		Hasta 10.000 habitantes	5,39	4,52	4,64	5,01		Hasta 10.000 habitantes	7,10	2,85	4,49	5,65		Hasta 10.000 habitantes	4,74	4,50	4,49	4,11			
		De 10.001 a 50.000 habitantes	4,27	5,02	4,64	5,70		De 10.001 a 50.000 habitantes	2,76	4,45	3,37	6,36		De 10.001 a 50.000 habitantes	3,31	5,09	4,08	6,20			
		De 50.001 a 100.000 habitantes	4,96	4,54	4,55	5,75		De 50.001 a 100.000 habitantes	3,14	3,67	1,17	2,33		De 50.001 a 100.000 habitantes	4,19	5,16	4,22	5,43			
		De 100.001 a 400.000 habitantes	5,97	5,44	5,05	5,91		De 100.001 a 400.000 habitantes	6,65	7,06	4,98	8,12		De 100.001 a 400.000 habitantes	6,29	4,28	5,63	5,64			
		Más de 400.000 habitantes	4,06	3,91	5,15	5,44		Más de 400.000 habitantes	7,70	5,06	5,99	12,57		Más de 400.000 habitantes	3,72	4,06	7,13	2,99			

Resultados

Tabla 23. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (45-64 años, más de 65 años).

			2006	2011	2014	2017				2006	2011	2014	2017
45-64 años	Ambos sexos	Total	3,67	3,15	3,53	3,89	65 y más años	Ambos sexos	Total	6,66	4,80	5,23	5,62
		Hasta 10.000 habitantes	3,81	2,61	3,06	3,23			Hasta 10.000 habitantes	7,08	5,35	5,63	6,43
		De 10.001 a 50.000 habitantes	3,67	3,72	3,89	3,48			De 10.001 a 50.000 habitantes	8,10	5,60	6,08	6,03
		De 50.001 a 100.000 habitantes	3,32	2,10	3,25	5,04			De 50.001 a 100.000 habitantes	5,32	3,67	6,64	5,38
		De 100.001 a 400.000 habitantes	3,31	3,62	4,10	4,02			De 100.001 a 400.000 habitantes	7,03	5,47	4,28	5,01
		Más de 400.000 habitantes	3,78	2,38	3,05	4,34			Más de 400.000 habitantes	4,53	3,04	3,87	4,52
	Varones	Total	3,08	2,01	2,67	2,68		Varones	Total	6,88	3,76	5,07	4,69
		Hasta 10.000 habitantes	3,59	2,23	2,67	2,53			Hasta 10.000 habitantes	7,35	3,75	5,02	5,89
		De 10.001 a 50.000 habitantes	3,36	3,26	2,69	2,41			De 10.001 a 50.000 habitantes	8,66	4,33	6,47	5,62
		De 50.001 a 100.000 habitantes	1,20	0,23	1,22	3,48			De 50.001 a 100.000 habitantes	4,88	2,28	7,79	5,14
		De 100.001 a 400.000 habitantes	2,21	1,43	3,76	2,91			De 100.001 a 400.000 habitantes	6,74	4,78	3,18	2,34
		Más de 400.000 habitantes	4,53	1,31	2,34	2,45			Más de 400.000 habitantes	5,20	2,77	3,79	3,65
	Mujeres	Total	4,23	4,25	4,38	5,08		Mujeres	Total	6,50	5,58	5,36	6,34
		Hasta 10.000 habitantes	4,03	3,05	3,51	4,00			Hasta 10.000 habitantes	6,88	6,66	6,17	6,89
		De 10.001 a 50.000 habitantes	3,99	4,16	5,03	4,52			De 10.001 a 50.000 habitantes	7,69	6,52	5,75	6,34
		De 50.001 a 100.000 habitantes	6,43	3,34	5,30	6,56			De 50.001 a 100.000 habitantes	5,71	4,74	5,85	6,64
		De 100.001 a 400.000 habitantes	4,38	5,75	4,40	5,00			De 100.001 a 400.000 habitantes	7,37	6,01	5,12	6,56
		Más de 400.000 habitantes	3,08	3,88	3,68	6,22			Más de 400.000 habitantes	4,09	3,23	3,93	5,20

En la serie temporal de promedio de todos los grupos de población, edad y sexo (tabla 24 y figuras 33 y 34), el grupo de mujeres tiene la cifra más alta en poblaciones de 100.000 a 400.000 habitantes (en todos los años). En 2014-2017 la prevalencia en grandes poblaciones de más de 400.000 se acerca a los otros grupos.

En varones la cifra más alta es en poblaciones de hasta 10.000 habitantes en 2006 y 2017. Se observan elevaciones en poblaciones de 100.000-400.000 habitantes y también en las más pequeñas de 10.000 pero no en poblaciones de más de 400.000 habitantes.

Resultados

Tabla 24. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 por tipos de población.

	Total																	
	Ambos sexos						Varones						Mujeres					
	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
2006	4,45	4,91	4,06	4,33	4,38	3,76	3,93	4,52	3,84	3,67	4,00	3,46	4,94	5,39	4,27	4,96	5,97	4,06
2011	4,08	4,17	4,56	3,45	4,50	3,18	3,38	3,83	4,06	2,34	3,50	2,38	4,75	4,52	5,02	4,54	5,44	3,91
2014	4,37	4,24	4,42	4,13	4,62	4,31	3,90	3,84	4,20	3,69	4,14	3,36	4,82	4,64	4,64	4,55	5,05	5,15
2017	4,68	4,58	4,81	4,81	4,64	4,54	3,74	4,16	3,86	3,82	3,23	3,59	5,57	5,01	5,70	5,75	5,91	5,44

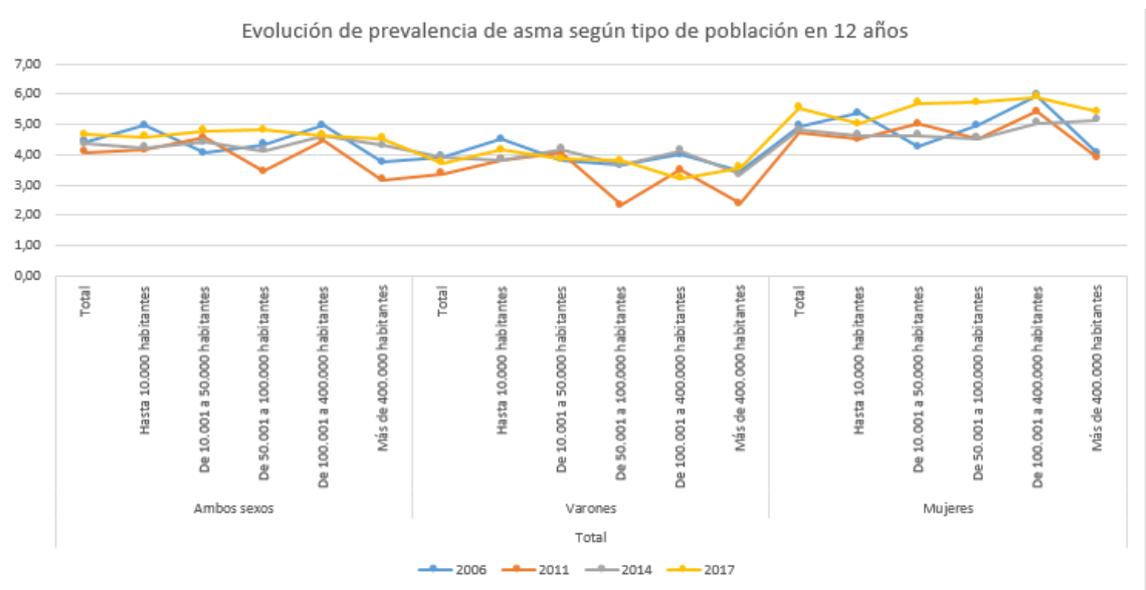


Figura 33. Promedio de la prevalencia del asma en 12 años (desglosados según color de la leyenda en 4 puntos (2006, 2011, 2014, 2017) por tipos de población incluidos en divisiones de cifras totales, varones y mujeres. El porcentaje de prevalencia representado en ordenadas.

Resultados

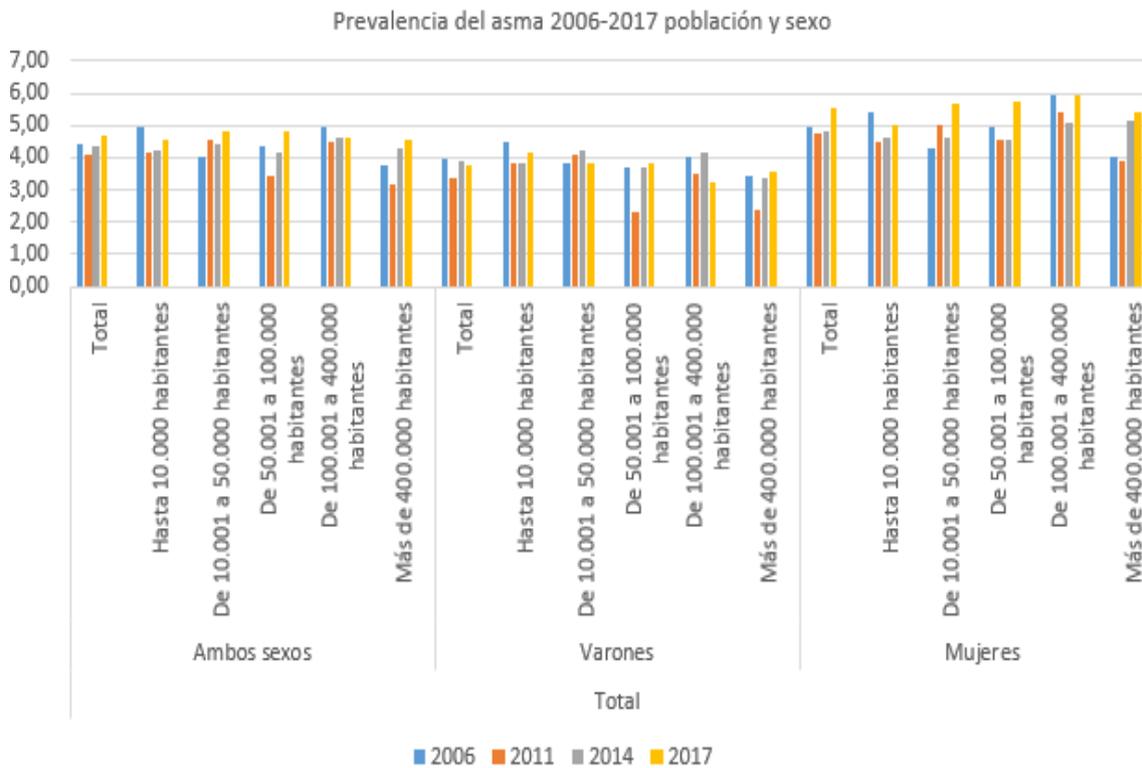


Figura 34. Diagrama de barras con el promedio del porcentaje de prevalencia en ordenadas y en abscisas los diferentes tipos de población con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

La estadística descriptiva del asma se refleja en la tabla 25 y en la figura 35. La media de prevalencia de asma, se mantiene en cifras totales, en poblaciones de 10.000 habitantes en 4,48%, mientras que en poblaciones de más de 400.000 habitantes es de 3,95%, guardando proporción similar en ambos sexos por separado. La mediana no muestra grandes diferencias con la media. La dispersión de la muestra es baja y ronda del 5 al 15%. La curtosis y la asimetría varían según tipo de población con distribuciones variables en un sentido y en el otro.

El test de Student no muestra significatividad, p -valor = 0,936 entre 2006 y 2017. La distribución sigue un patrón normal (Shapiro-Wilk $W = 0,96926$, p -valor = 0,6488).

Resultados

Tabla 25. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a tipo de población y sexo.

	Total																	
	Ambos sexos					Varones					Mujeres							
	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	Más de 100.001 a 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	Más de 100.001 a 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	Más de 100.001 a 400.000 habitantes			
Media	4,3925934	4,4899221	4,461459	4,1798222	4,6854059	3,9500351	3,7362623	4,0864097	3,9877424	3,3803552	3,7196193	3,1961606	5,0182625	4,8914829	4,9094715	4,9491508	5,5926942	4,638938
Error típico	0,123546	0,1844911	0,1564106	0,2824439	0,1018283	0,3033717	0,1271715	0,162876	0,0863563	0,3487195	0,2131433	0,2762164	0,1871555	0,1958691	0,3047063	0,2831716	0,216087	0,3857771
Mediana	4,408059	4,4092322	4,4897031	4,2305373	4,6327944	4,0383456	3,8189376	4,0020867	3,9593655	3,679897	0,2131433	3,4085414	4,8800008	4,8290078	4,8324907	4,7570185	5,6735481	4,6038768
Desviación estándar	0,247092	0,3689821	0,3128212	0,5648878	0,2036566	0,6067433	0,2543431	0,325752	0,1727126	0,697439	0,4262866	0,5524328	0,374311	0,3917381	0,6094126	0,5663432	0,432174	0,7715541
Varianza de la muestra	0,0610545	0,1361478	0,0978571	0,3190982	0,041476	0,3681375	0,0646904	0,1061144	0,0298296	0,4864212	0,1817203	0,305182	0,1401087	0,1534588	0,3713837	0,3207446	0,1867743	0,5952958
Curiosis	1,1068369	-1,141139	0,7448498	1,014888	2,6082476	-1,525221	1,576055	-0,992137	-3,162679	3,8240862	-3,66016	3,4568872	3,0364091	-1,643573	-0,005861	1,2922231	-2,34097	-5,135493
Coefficiente de asimetría	-0,361537	0,839187	-0,501305	-0,508982	1,4065894	-0,609332	-1,414502	0,9086517	0,4824608	-1,942209	-0,242229	-1,827955	1,7303257	0,6417641	0,6412929	1,369504	-0,6249	0,0880485
Rango	0,5987806	0,8022663	0,7496963	1,3607102	0,47568	1,3578709	0,5545117	0,6900331	0,3606444	1,483306	0,9110981	1,208577	0,8197726	0,8680595	1,4258884	1,209121	0,9198383	1,5354767
Mínimo	4,0777374	4,1694789	4,0583667	3,4487519	4,5001773	3,1827891	3,3763311	3,8257162	3,8357971	2,3391604	3,231623	2,3794912	4,7466378	4,5199283	4,273508	4,5367227	5,0519212	3,9062609
Máximo	4,676518	4,9717452	4,808063	4,8094621	4,9758573	4,5406601	3,9308428	4,5157493	4,1964415	3,8224664	4,142721	3,5880682	5,5664105	5,3879878	5,6993964	5,7458437	5,9717595	5,4417376
Suma	17,570373	17,959688	17,845836	16,719289	18,741623	15,80014	14,945049	16,345639	15,95097	13,521421	14,878477	12,784642	20,07305	19,565932	19,637886	19,796603	22,370777	18,555752

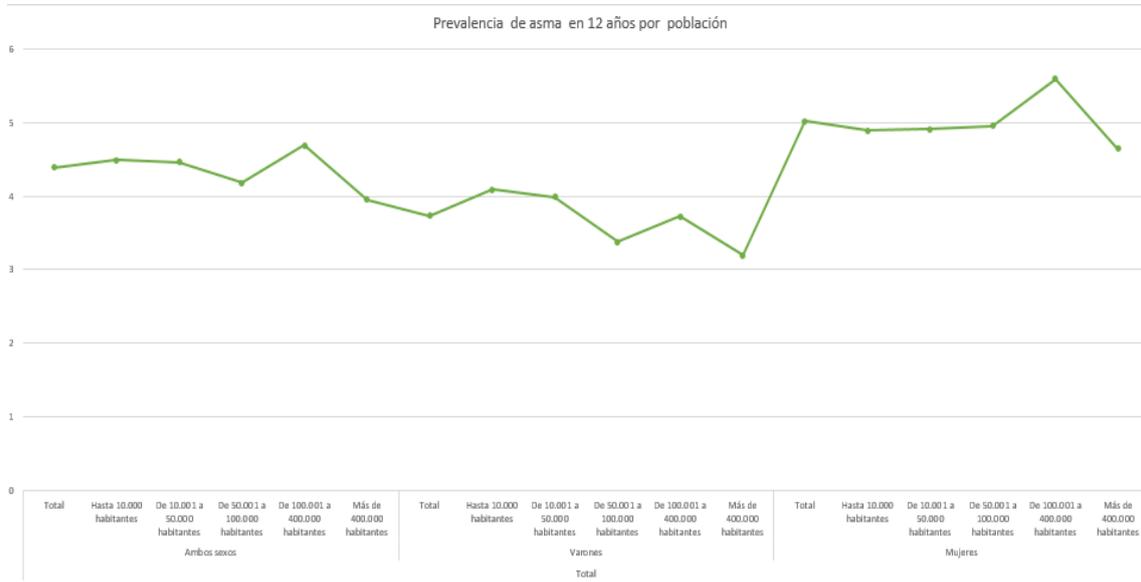


Figura 35. Promedio de la prevalencia del asma en 12 años por tipo de población incluidos en divisiones de cifras totales, varones y mujeres.

Resultados

Tabla 26. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y los tipos de población y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Total																	
		Ambos sexos					Varones					Mujeres							
Total		Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	

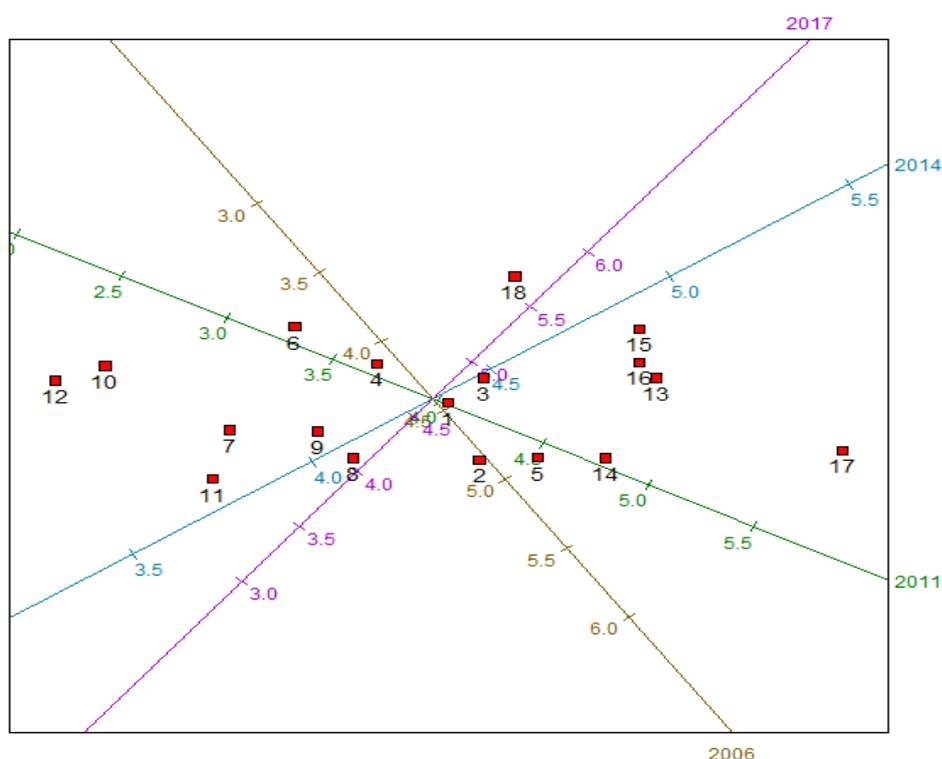


Figura 36. Biplot de representación de evolución de asma según tipo de población.

La figura 36 representa a las variables que son cada uno de los grupos de edad y sexo, situándose sobre los ejes el momento temporal en que se efectúan las observaciones, con la divisiones representado las cifras de prevalencia. La tabla 26 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 36 y que corresponden a diferentes nichos de edad y sexo.

Resultados

Se presenta la evolución de la prevalencia de asma en los años indicados en ejes según tipo de población. Ocupando las posiciones más extremas con proyección a valores bajos de prevalencia, los grupos 10 que corresponde a varones de poblaciones de 50.000-100.000 habitantes y 12 a varones de poblaciones de más de 400.000 habitantes. Sin embargo, también en posición periférica pero con proyección a valores altos de prevalencia tenemos a los nichos 15, 16, 17 y 18; todos ellos correspondiente a mujeres pero con independencia del tipo de núcleo de población. Teniendo en cuenta el ángulo que forman los ejes, la posición ortogonal entre 2006 y 2017 nos habla de escasa correlación.

4.1.9. Estadística descriptiva de la prevalencia de asma según nivel de estudios

Si tenemos en cuenta el nivel de estudios clasificándolo en: sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado; en la tabla 27 vemos, que en cifras totales, el grupo de sin estudios tiene prevalencias más altas en asma: 7,80% en 2006, 6,22% en 2011 y 7,33% en 2017. Sobre todo con respecto al grupo de tercer grado que oscila entre 3,33% en 2006, 3,55% en 2011, 3,88% en 2014 y 4,63% en 2017. Además, en varones sin estudios las cifras son más altas, de casi el 9% en 2006. En mujeres sin estudios el aumento es de un 30% más. Hay un marcado aumento en la población sin estudios, tanto de ambos sexos como por separado.

Tabla 27. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 por nivel de estudios.

	Total																	
	Ambos sexos						Varones						Mujeres					
	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
2006	4,45	7,80	4,61	4,43	3,64	3,33	3,93	8,63	4,33	3,88	2,54	3,16	4,94	7,26	4,85	5,00	4,81	3,51
2011	4,08	6,22	4,12	3,87	3,73	3,55	3,38	4,42	4,47	3,40	2,89	2,80	4,75	7,50	3,73	4,36	4,59	4,21
2014	4,37	5,71	4,58	4,28	4,11	3,88	3,90	5,26	4,35	3,84	3,46	3,56	4,82	6,00	4,80	4,80	4,80	4,14
2017	4,68	7,33	4,70	4,19	4,23	4,63	3,74	6,37	3,91	3,14	3,09	4,47	5,57	7,95	5,42	5,33	5,40	4,75

Resultados

Tomando como puntos de medida el inicio y el final de la serie temporal, el test de Student no muestra significatividad p -valor = 0,936 entre 2006 y 2017. La distribución sigue un patrón normal (Shapiro-Wilk $W = 0,98057$, p -valor = 0,1267).

Las figuras 37 y 38 reflejan claramente estos datos. Los picos máximos de prevalencia corresponden a los grupos con menor nivel educativo. La situación se mantiene tanto si consideramos cifras en ambos sexos como por separado en varones y mujeres.

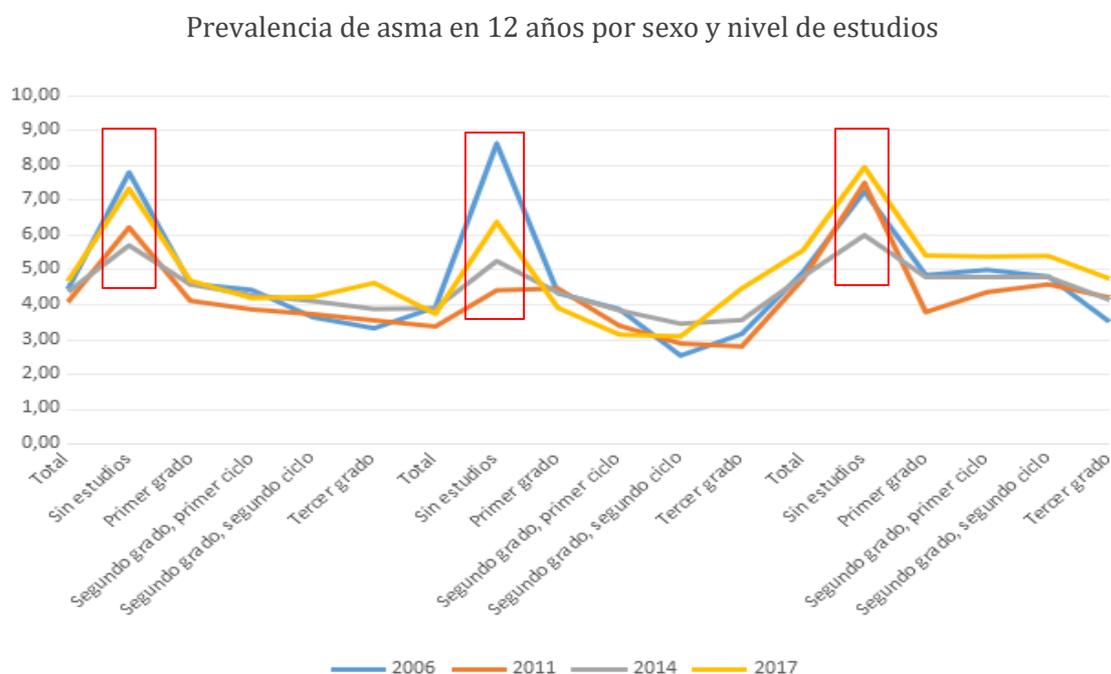


Figura 37. Evolución de la prevalencia de asma en 12 años, desglosados según color de la leyenda en 4 puntos (2006, 2011, 2014, 2017) de realización de encuesta, con el porcentaje de prevalencia representado en ordenadas y en abscisas cada uno de los grupos de nivel de estudios y sexo.

Resultados

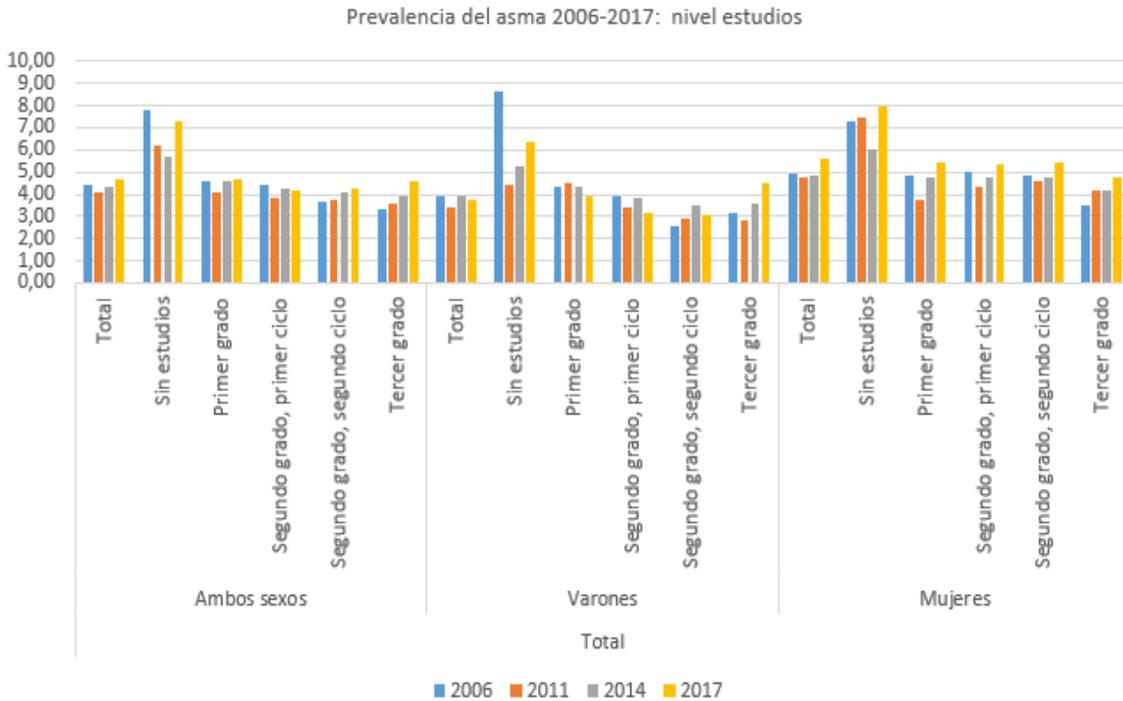


Figura 38. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de niveles de estudios con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Tabla 28. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a nivel de estudios y sexo.

	Ambos sexos						Varones						Mujeres					
	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
Media	4,3925934	6,7667398	4,5007981	4,1925666	3,9270353	3,8464835	3,7362623	6,1698668	4,2654429	3,5651322	2,9940685	3,4998641	5,0182625	7,1778291	4,7140623	4,887524	4,9013783	4,1529161
Error típico	0,123546	0,4842946	0,1306497	0,1184711	0,1417528	0,2838579	0,1271715	0,9130114	0,1222173	0,1776448	0,1915639	0,3589015	0,1871555	0,4185033	0,3383005	0,2128694	0,1755875	0,2543884
Mediana	4,408059	6,7785837	4,5959255	4,2357411	3,9190187	3,7167382	3,8189376	5,8154286	4,340401	3,6199912	2,9908195	3,3624366	4,8800008	7,3787865	4,8255758	4,9003567	4,8066278	4,1756446
Desviación e	0,247092	0,9685893	0,2612995	0,2369421	0,2835056	0,5677158	0,2543431	1,8260228	0,2444346	0,3552896	0,3831278	0,717803	0,374311	0,8370065	0,6766009	0,4257389	0,351175	0,5087768
Varianza de l	0,0610545	0,9381652	0,0682774	0,0561416	0,0803754	0,3223012	0,0646904	3,3343593	0,0597483	0,1262307	0,1467869	0,5152412	0,1401087	0,7005799	0,4577888	0,1812536	0,1233239	0,2588538
Curtosis	1,1068369	-3,635668	3,4163678	1,4603122	-4,7006001	0,974107	1,576055	0,5405345	3,0080783	-3,512047	0,1325972	0,7848121	3,0364091	2,1424889	1,9803128	0,3876069	2,734311	1,4544158
Coefficiente d	-0,361537	-0,041142	-1,792707	-0,993315	0,0697453	1,1230063	-1,414502	0,9612379	-1,606586	-0,423146	0,0459986	0,9738534	1,7303257	-1,28996	-0,950779	-0,166428	1,4605729	-0,265791
Rango	0,5987806	2,0957526	0,5802963	0,5603492	0,5830895	1,3005529	0,5545117	4,2138757	0,5597144	0,7345879	0,9156551	1,6675087	0,8197726	1,9550559	1,6260596	1,0215147	0,8172259	1,2427983
Mínimo	4,0777374	5,7070196	4,1155226	3,8692175	3,6435072	3,3259523	3,3763311	4,4173672	3,9106275	3,1429793	2,53949	2,8035371	4,7466378	5,9993438	3,7895191	4,363934	4,5875159	3,5087883
Máximo	4,676518	7,8027722	4,6958189	4,4295666	4,2265967	4,6265052	3,9308428	8,6312429	4,470342	3,8775672	3,455145	4,4710458	5,5664105	7,9543997	5,4155787	5,3854487	5,4047417	4,7515867
Suma	17,570373	27,066959	18,003192	16,770266	15,708141	15,385934	14,945049	24,679467	17,061772	14,260529	11,976274	13,999456	20,07305	28,711316	18,856249	19,550096	19,605513	16,611664

Resultados

La tabla 28 nos muestra la estadística descriptiva de la prevalencia de asma teniendo en cuenta los niveles de formación. Los grupos con niveles educativos bajos tienen medias más altas en prevalencia de asma que los grupos con niveles educativos más altos.

Estas diferencias no solo eran en cifras absolutas sino que se tradujeron en significatividad estadística. El test de Barlett sugería desigualdad en las varianzas, $F=0,12033$, $p\text{-valor} = 0,004192$ (se rechazaba la hipótesis nula de homocedasticidad). Por ello, practicamos el test de Welch comparando las medias de los niveles de estudios (los más bajos con los más elevados) en sus cifras de prevalencia (las menores frente a las mayores) apareciendo significatividad ($p\text{-valor} = 0,000575$).

Las desviaciones estándar están entre el 5-20% de los valores de la media. El apuntamiento de la curtosis es hacia la derecha y la asimetría se reparte a ambos lados según el grupo analizado.



Figura 39. Promedio de la prevalencia del asma en 12 años por nivel de estudios incluidos en divisiones de cifras totales, varones y mujeres.

Resultados

Tabla 29. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y los niveles de estudios y numerando cada una de estas divisiones.

Total																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
15-24 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
25-44 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
45-64 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
65 y más años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado

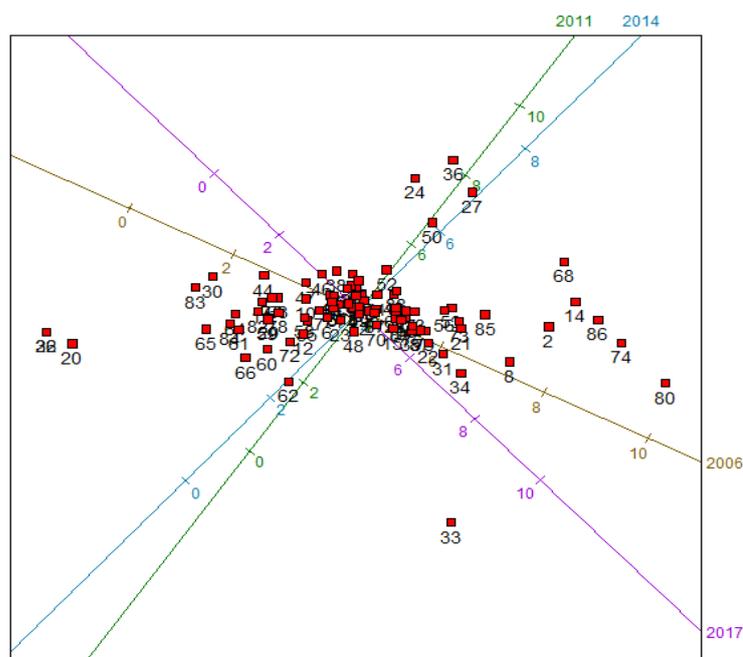


Figura 40. Biplot de representación de evolución de asma según nivel de estudios.

La figura 40 representa la evolución de la prevalencia de asma en los años indicados en los ejes según el nivel de estudios. Los ejes son el tiempo en que se efectúan las observaciones con la divisiones representado las cifras de prevalencia. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia y sobre ellos cada uno de los nichos de variables de edad y sexo.

La tabla 29 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 40 y que corresponden a diferentes nichos de edad y sexo.

Los grupos 80, 74 y 86 representan a los de más de 65 años sin estudios, es decir los de mayor prevalencia, situándose claramente separados en la zona de prevalencia más elevada. Los grupos de los nichos 20 de 15-24 años sin estudios de ambos sexos y 26 de 15-24 años sin estudios varones se proyectan en zonas de prevalencias menores. El nicho 33 de mujeres de 15-24 años con 1º grado, aparece en zona aislada.

La proximidad y mezcla entre puntos de observación se interpreta en términos de similitud en cuanto a cifras de prevalencia. Sobre la correlación entre ejes teniendo en cuenta el ángulo, destacamos la independencia de 2017 frente a 2011-2014.

4.1.10. Estadística descriptiva de la prevalencia de asma según nivel económico

Para caracterizar el nivel socio-económico hemos usado la clase social ocupacional, la cual obtiene la clase social del individuo que trabaja o ha trabajado y del sustentador principal y se escoge la más privilegiada de estas dos, agrupando en las clases sociales I y II (más privilegiadas), III, IV, V y VI (menos privilegiadas) de acuerdo con la adaptación utilizada por la Encuesta Nacional de Salud (218, 219).

La tabla 30 nos muestra los registros de prevalencia del asma teniendo en cuenta cuatro puntos de medición en la serie temporal: 2006, 2011, 2014 y 2017.

Los niveles económicos más altos muestran valores bajos de asma: en nivel I 2,77% en 2006, 2,93% en 2011, 3,46% en 2014 y 4% en 2017.

En el nivel económico más bajo que es el VI: 4,76% en 2006, 4,94% en 2011, 4,67% tanto en 2014 como en 2017.

Resultados

Tabla 30. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 según nivel económico.

	Total																				
	Ambos sexos						Varones						Mujeres								
	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
2006	4,45	2,77	3,89	4,05	4,77	5,75	4,76	3,93	3,13	3,27	3,35	4,61	4,77	3,80	4,94	2,37	4,52	4,74	4,94	6,55	5,61
2011	4,08	2,93	3,51	3,35	3,89	4,59	4,94	3,38	2,08	2,85	2,53	3,61	3,87	4,15	4,75	3,88	4,08	4,08	4,22	5,31	5,55
2014	4,37	3,46	4,03	4,76	4,41	4,38	4,67	3,90	3,45	3,99	4,26	4,02	3,93	3,62	4,82	3,47	4,07	5,24	4,82	4,93	5,56
2017	4,68	4,00	4,25	4,22	4,91	5,08	4,67	3,74	4,13	3,19	3,22	3,43	4,14	3,87	5,57	3,88	5,25	5,18	6,66	5,99	5,33

El test de Welch, entre las medias de prevalencia del punto inicial de medición 2006 y el punto final 2017, no muestra significatividad, p-valor = 0,5932. La distribución no sigue un patrón normal (test de Shapiro-Wilk $W = 0,96716$, p-valor = 0,0224, $p < 0,05$ se rechaza la hipótesis nula de normalidad).

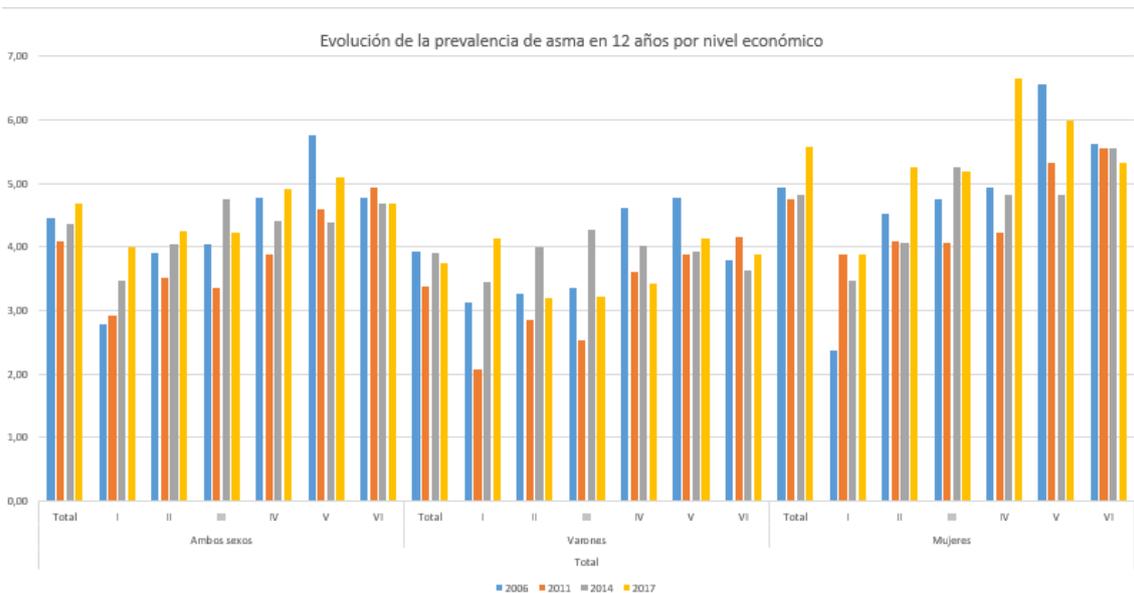


Figura 41. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes niveles económicos con 3 divisiones: ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Resultados

La figura 41 nos expone, en forma de diagrama de barras, los registros conseguidos en la tabla 30. Sitúa el porcentaje de prevalencia en ordenadas y en abcisas los diferentes niveles económicos con 3 divisiones: ambos sexos, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

Tabla 31. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto al nivel económico incluyendo divisiones de cifras totales, varones y mujeres.

	Total																				
	Ambos sexos						Varones						Mujeres								
	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
Media	4,392593	3,291267	3,919927	4,095664	4,495349	4,950836	4,763319	3,736262	3,198638	3,324171	3,339009	3,917382	4,177568	3,86051	5,018262	3,39999	4,479521	4,807711	5,161088	5,669497	5,514375
Error típico	0,123546	0,279404	0,156355	0,291301	0,228877	0,304323	0,063446	0,127172	0,426932	0,239763	0,356781	0,262149	0,204744	0,109645	0,187156	0,3565	0,275982	0,272017	0,524469	0,378906	0,062057
Mediana	4,408059	3,194481	3,961708	4,137067	4,592124	4,834618	4,718588	3,818938	3,290146	3,228798	3,284474	3,815957	4,034261	3,834276	4,880001	3,674506	4,301525	4,962702	4,883788	5,649474	5,556669
Desviación estándar	0,247092	0,558807	0,31271	0,582601	0,457754	0,608647	0,126891	0,254343	0,853864	0,479525	0,713561	0,524297	0,409487	0,219291	0,374311	0,712999	0,551965	0,544034	1,048937	0,757812	0,124114
Varianza de la muestr	0,061054	0,312266	0,097788	0,339424	0,209539	0,370451	0,016101	0,06469	0,729084	0,229944	0,50917	0,274888	0,16768	0,048089	0,140109	0,508368	0,304665	0,295973	1,10027	0,574279	0,015404
Curtosis	1,106837	-1,57344	0,919619	1,16844	-0,33927	-0,63879	1,511785	1,576055	1,085169	2,104291	1,455999	-0,32434	2,380529	1,299104	3,036409	2,388514	0,779649	0,307446	2,644156	-1,92205	3,277154
Coefficiente de asimet	-0,36154	0,686674	-0,73019	-0,41098	-0,92398	0,833926	1,416678	-1,4145	-0,6047	1,119924	0,450053	0,874095	1,586517	0,680315	1,730326	-1,60735	1,263205	-1,15298	1,431562	0,116496	-1,72709
Rango	0,598781	1,232207	0,74404	1,4117	1,024231	1,366115	0,269545	0,554512	2,051992	1,140542	1,73631	1,185137	0,894967	0,527493	0,819773	1,511565	1,176428	1,179298	2,443829	1,72837	0,279453
Mínimo	4,077737	2,771949	3,506127	3,348412	3,886458	4,383997	4,673277	3,376331	2,081135	2,849274	2,52539	3,426239	3,873391	3,622997	4,746638	2,369692	4,069302	4,063072	4,216472	4,825334	5,332354
Máximo	4,676518	4,004156	4,250167	4,760112	4,910689	5,750112	4,942822	3,930843	4,133127	3,989816	4,2617	4,611376	4,768358	4,15049	5,56641	3,881257	5,24573	5,24237	6,660301	6,553704	5,611807
Suma	17,57037	13,16507	15,67971	16,38266	17,98139	19,80334	19,05327	14,94505	12,79455	13,29669	13,35604	15,66953	16,71027	15,44204	20,07305	13,59996	17,91808	19,23085	20,64435	22,67799	22,0575

La tabla 31 expone los valores de estadística descriptiva de la prevalencia de asma teniendo en cuenta los niveles de estudio.

Los niveles económicos más altos muestran medias con valores bajos de asma, y, al contrario, en los niveles económicos más bajos se elevan las cifras.

La mediana no muestra grandes diferencias con la media. La dispersión de la muestra es baja y ronda del 3 al 15%. El apuntamiento de la curtosis es mayoritariamente a la derecha y la asimetría varía según grupo con distribuciones variables en un sentido y en el otro.

La diferencia de prevalencia, entre las clases socio-económicas más bajas y las más altas, tiene una traducción a nivel de significatividad estadística. Dado que el test de Barlett sugería homocedasticidad ($F = 0,83587$, $p\text{-valor} = 0,6709$), aplicamos el test de Student con significatividad muy clara con $p\text{-valor} = 0,00082$.

Resultados



Figura 42. Promedio de la prevalencia del asma en 12 años según nivel económico.

La figura 42 nos permite dibujar las medias de prevalencia del asma según nivel económico. Se refleja claramente el concepto de valores bajos de asma en niveles altos y, al contrario, en los niveles económicos más bajos presentan medias altas de asma. En mujeres es claro, ya que el máximo roza un 6% en clases bajas siendo la mitad en clases altas.

Resultados

Tabla 32. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el nivel económico y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Total																				
Ambos sexos							Varones							Mujeres						
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
15-24 años																				
Ambos sexos							Varones							Mujeres						
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
25-44 años																				
Ambos sexos							Varones							Mujeres						
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84
45-64 años																				
Ambos sexos							Varones							Mujeres						
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI
85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
65 y más años																				
Ambos sexos							Varones							Mujeres						
Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI	Total	I	II	III	IV	V	VI

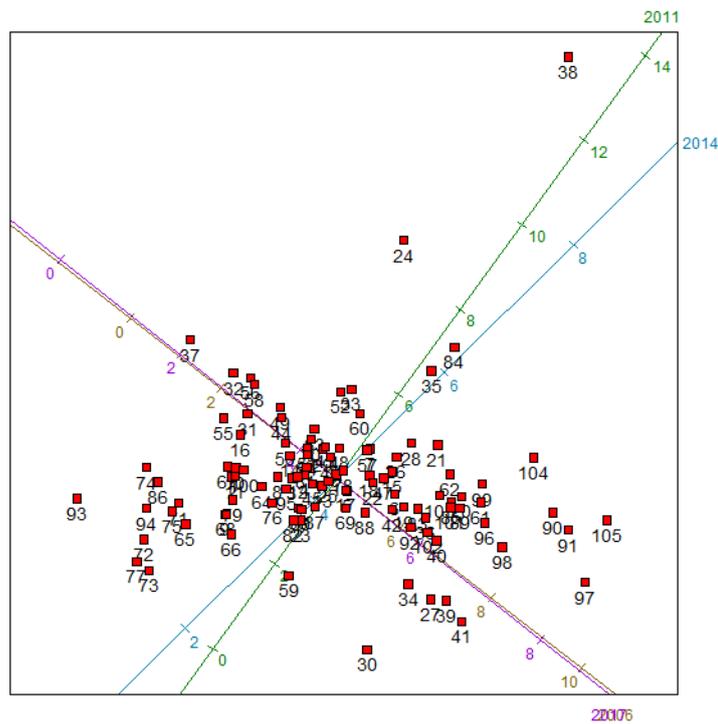


Figura 43. Biplot de representación de evolución de asma según nivel económico.

La figura 43 muestra un biplot de evolución de prevalencia del asma según nivel económico. Representa sobre los ejes a las variables que son cada uno de los grupos de niveles económico, edad y sexo. Cada uno de los ejes son el momento temporal, situándose sobre los ejes los porcentajes de prevalencia.

La tabla 32 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 43 y que corresponden a diferentes nichos de nivel económico edad y sexo.

Se representa la evolución de la prevalencia de asma en los años indicados en ejes según el nivel económico. Ocupando las posiciones más extremas: nicho 24 de 15-24 años de ambos sexos del nivel II y nicho 38 de 15-24 años mujeres del nivel II. Situándose de forma muy próxima entre puntos, en la zona de los ejes con elevadas prevalencias, los nichos correspondientes, y por el contrario, en las zonas de ejes de bajas prevalencias los nichos que responden a las clases menos favorecidas.

Poca correlación angular entre 2017 y 2011-2014.

Resultados

4.1.1.1. Estadística descriptiva de la prevalencia de asma según tipo de ocupación

No aparecen variaciones que sean de importancia si comparamos la evolución de la prevalencia de asma respecto a la situación laboral de los sujetos.

La tabla 33 nos ofrece las cifras de prevalencia de asma por ocupación y desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (15-24 años, 25-44 años, 45-64 años, más de 65 años). Lógicamente las cifras de prevalencia son más irregulares porque algunos grupos bien por juventud o por jubilación no están cubiertos.

Tabla 33. Cifras de prevalencia de asma por ocupación, en los cuatro años de medición, y desglosadas por totales de ambos sexos, varones y mujeres distribuidos en grupos de edad (15-24 años, 25-44 años, 45-64 años, más de 65 años).

Total											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
4,45	3,39	5,32	5,64	3,93	2,96	4,20	5,87	4,94	4,04	6,19	5,51
4,08	3,34	4,50	4,75	3,37	2,89	3,73	3,95	4,75	3,88	5,50	5,26
4,34	3,87	4,49	4,85	3,80	3,30	3,24	4,90	4,86	4,59	5,73	4,81
4,66	4,16	4,62	5,25	3,79	3,57	3,16	4,34	5,49	4,87	5,99	5,31
16-24 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
5,03	5,25	6,05	4,72	4,72	4,25	6,70	4,83	5,35	6,60	5,55	4,62
4,76	5,75	4,95	4,51	4,70	5,09	4,88	4,61	4,81	6,38	5,07	4,40
3,61	4,87	4,09	3,16	2,85	3,33	3,46	2,53	4,42	7,19	4,88	3,73
5,16	5,85	4,40	5,12	3,85	4,70	2,92	3,83	6,50	6,99	6,06	6,46
25-44 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
3,74	3,23	6,40	4,79	3,07	2,92	4,36	3,90	4,45	3,67	7,79	4,98
4,25	3,77	5,28	5,21	3,88	3,64	4,66	4,30	4,63	3,93	6,02	5,49
4,75	4,49	5,12	5,74	4,42	4,03	4,15	11,52	5,08	5,05	5,96	3,55
4,66	4,78	4,84	3,38	4,34	4,55	3,98	2,35	4,98	5,04	5,58	3,75
45-64 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
3,67	3,14	2,81	4,65	3,08	2,62	2,70	4,74	4,23	4,01	2,93	4,61
3,15	2,48	2,94	4,49	2,01	1,68	1,68	3,30	4,25	3,54	4,65	4,98
3,53	2,93	3,60	4,92	2,67	2,37	1,97	4,80	4,38	3,66	5,43	4,97
3,89	3,22	4,41	5,27	2,68	2,30	2,27	4,58	5,08	4,37	6,48	5,63
65 y más años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
6,66	3,88	0	6,73	6,88	6,17	0	6,31	6,50	0	0	6,59
4,80	0,67	0	4,86	3,76	0,00	0	3,84	5,58	1,33	0	5,63
5,23	4,12	0	5,24	5,07	1,95	0	5,14	5,36	6,97	0	5,31
5,62	6,97	0	5,59	4,69	5,61	0	4,66	6,34	8,21	0	6,31

Resultados

La tabla 34 muestra los valores de la estadística descriptiva de la prevalencia de asma en relación con el tipo de ocupación. El test de Welch, sobre las medias del inicio de la medición con respecto a la última, no muestra significación estadística (p -valor= 0,6992) y la distribución no sigue una distribución normal (Shapiro-Wilk, $W = 0,93814$, p -valor = 0,00443). El único dato de interés es que el punto más alto lo da el grupo de mujeres en paro.

Tabla 34. Estadística descriptiva de la evolución de la prevalencia de asma en 12 años con respecto a ocupación y sexo.

	Total											
	Ambos sexos				Varones				Mujeres			
	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
Media	4,3817691	3,6901529	4,7343023	5,1230597	3,721249	3,1790389	3,5818175	4,764663	5,0108608	4,3447621	5,852598	5,3712095
Error típico	0,1214024	0,1953139	0,1984821	0,2048576	0,1210995	0,1571263	0,239732	0,4173204	0,1649468	0,2311676	0,1514894	0,229642
Mediana	4,3933192	3,6312029	4,5617783	5,05066	3,7912413	3,1259951	3,4833615	4,6198809	4,9003278	4,3136187	5,8584506	5,3848421
Desviación estándar	0,2428048	0,3906278	0,3969642	0,4097152	0,2421991	0,3142526	0,4794641	0,8346408	0,3298936	0,4623352	0,3029788	0,459284
Varianza de la muestra	0,0589542	0,1525901	0,1575805	0,1678666	0,0586604	0,0987547	0,2298858	0,6966252	0,1088298	0,2137538	0,0917962	0,2109418
Curtosis	0,6230956	-3,3910273	3,487012	-1,8259382	2,8838945	-2,4141024	-1,6619715	0,0195347	2,9649789	-3,8450687	-1,7813261	0,1235971
Coefficiente de asimetría	-0,2663777	0,4343301	1,8645278	0,667987	-1,5398835	0,5834932	0,7455402	0,8381884	1,6653283	0,2073773	-0,0867722	-0,1604604
Rango	0,5842897	0,8145253	0,8316877	0,8943803	0,5591721	0,6750552	1,0335382	1,9206959	0,7403223	0,9830628	0,6938483	1,0964727
Mínimo	4,0780743	3,3418402	4,4909824	4,7482693	3,3716707	2,894555	3,1635043	3,9490971	4,7512328	3,8843741	5,4998211	4,8093406
Máximo	4,662364	4,1563655	5,32267	5,6426496	3,9308428	3,5696102	4,1970425	5,869793	5,491555	4,8674369	6,1936695	5,9058133
	17,527077	14,760611	18,937209	20,492239	14,884996	12,716156	14,32727	19,058652	20,043443	17,379048	23,410392	21,484838

Las desviaciones estándar están entre el 5-20% de los valores de la media. El apuntamiento de la curtosis y la asimetría se reparten a ambos lados según el grupo analizado.

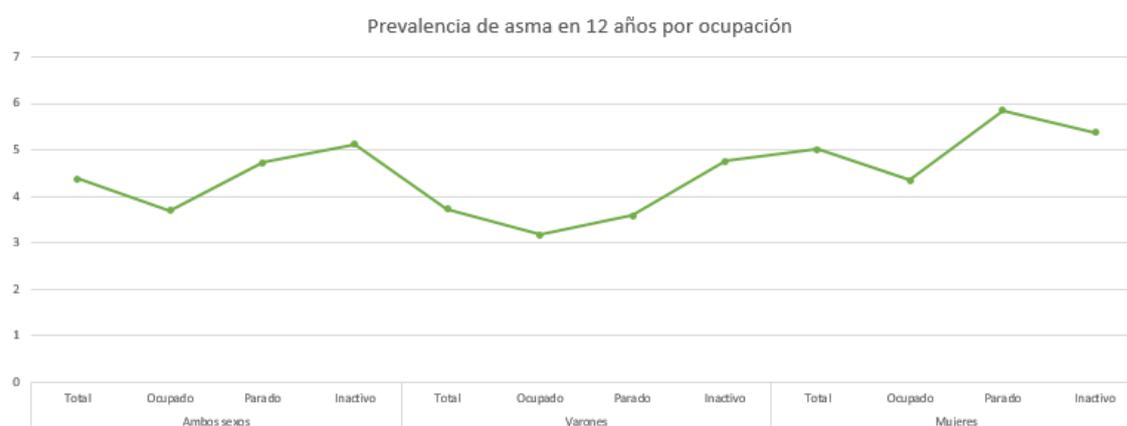


Figura 44. Promedio de la prevalencia del asma en 12 años por tipos de ocupación incluidos en divisiones de cifras totales, varones y mujeres.

Resultados

Como nos señala la figura 44, la cual grafica las cifras de media de la estadística descriptiva, como media en mujeres paradas es evidente que se alcanza el máximo valor. Los inactivos muestran más prevalencia de asma que los ocupados, tanto en cifras de ambos sexos como por separado en varones y mujeres.

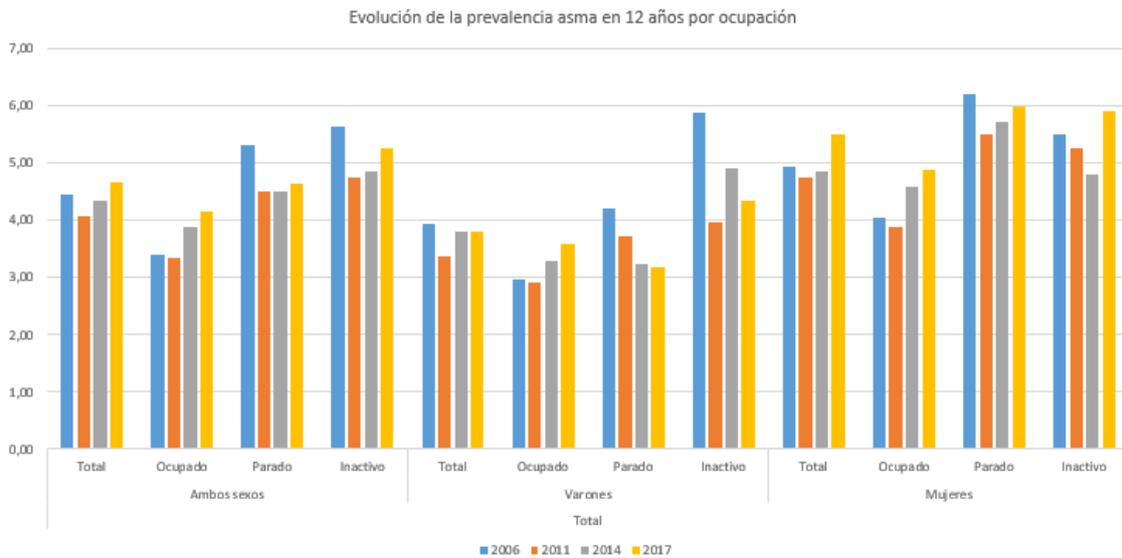


Figura 45. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abcisas los diferentes tipos de ocupación con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realiza la encuesta (relación de año con su color en la leyenda del pie de la figura).

En la figura 45, en forma de diagrama de barras, vemos que los inactivos, tanto en cifras totales como en ambos sexos por separado, muestran una prevalencia más alta que los ocupados. En mujeres las cifras se equiparan entre las inactivas y las paradas. Lógicamente esto es así porque la población inactiva es la de mayor edad y en ellos aparece un claro aumento del asma como hemos visto en anteriores apartados.

Resultados

Tabla 35. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el tipo de ocupación y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.

1	2	3	4	5	6	7	8	9	10	11	12
Total											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
13	14	15	16	17	18	19	20	21	22	23	24
16-24 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
25	26	27	28	29	30	31	32	33	34	35	36
25-44 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
37	38	39	40	41	42	43	44	45	46	47	48
45-64 años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo
49	50	51	52	53	54	55	56	57	58	59	60
65 y más años											
Ambos sexos				Varones				Mujeres			
Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo	Total	Ocupado	Parado	Inactivo

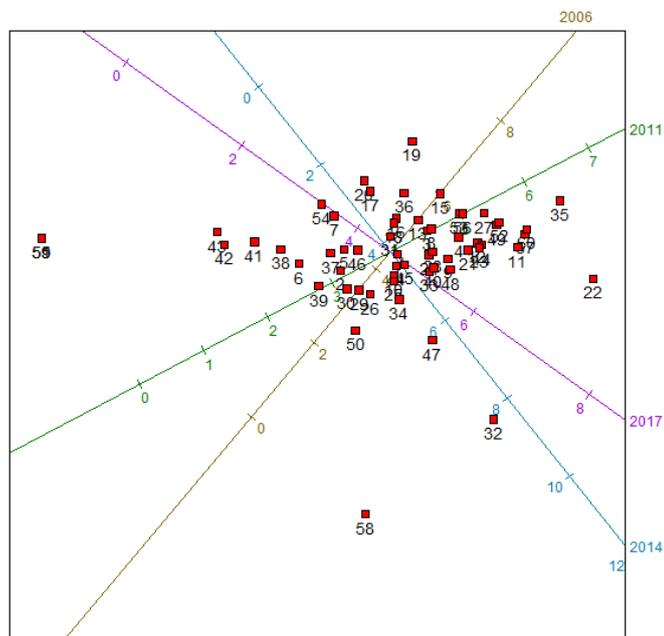


Figura 46. Biplot de representación de evolución de asma según ocupación.

Resultados

La figura 46 representa la evolución de la prevalencia de asma en los años indicados en ejes según el tipo de ocupación. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia, y sobre los ejes las variables que son cada uno de los grupos de tipo de ocupación, sexo y edad. La tabla 35 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 46 y que corresponden a diferentes nichos de ocupación, edad y sexo.

Los puntos con menos relación son lógicamente los del grupo de mayor edad de más de 65 años. También el 22 (mujeres ocupadas de 15-24 años) y el 32 (varones inactivos de 25-44 años).

Podemos observar que de 2011 a 2017 hay una correlación entre los datos por el ángulo que mantiene entre ellos, siendo la correlación nula entre 2014 y 2006.

4.1.12. Estadística descriptiva de la prevalencia de asma según tipo de jornada laboral

La tabla 36 y la figura 47 nos muestran los registros de prevalencia del asma teniendo en cuenta cuatro puntos de medición en la serie temporal: 2006, 2011, 2014 y 2017, y anotando que tipo de jornada laboral tenían los participantes.

Aparecen variaciones y dispersión de la muestra dadas las grandes diferencias en la jornada laboral que pueden producirse en el tiempo. Así sorprende que haya puntos más altos, en cifras totales de ambos sexos, en jornada reducida con respecto a la jornada partida.

Tabla 36. Cifras de prevalencia de asma desglosadas por totales de ambos sexos, varones y mujeres distribuidos entre 2006 y 2017 según tipo de jornada laboral.

	Total																		
	Ambos sexos						Varones						Mujeres						
	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	
2006	3,39	2,64	4,59	3,17	2,88	3,17	2,94	2,64	3,78	1,23	2,41	2,82	4,05	3,28	5,38	2,68	3,63	3,66	
2011	3,36	3,10	3,29	5,06	2,94	3,92	2,91	2,71	2,80	7,24	3,16	2,94	3,90	3,75	3,71	4,60	2,65	5,28	
2014	3,87	3,48	4,32	6,42	3,71	3,40	3,31	3,38	3,30	5,25	3,32	2,76	4,57	3,66	5,25	6,75	4,30	4,27	
2017	4,18	3,74	4,59	4,02	3,65	4,48	3,55	3,44	3,73	1,02	3,36	3,85	4,83	4,23	5,38	4,84	4,03	5,24	

Resultados

El test de Welch sobre las medias del inicio de la medición con respecto a la última, no muestra significación estadística (p-valor = 0,7765) y la distribución no sigue un patrón normal (Shapiro-Wilk $W = 0,92453$, p-valor = 0,00034).

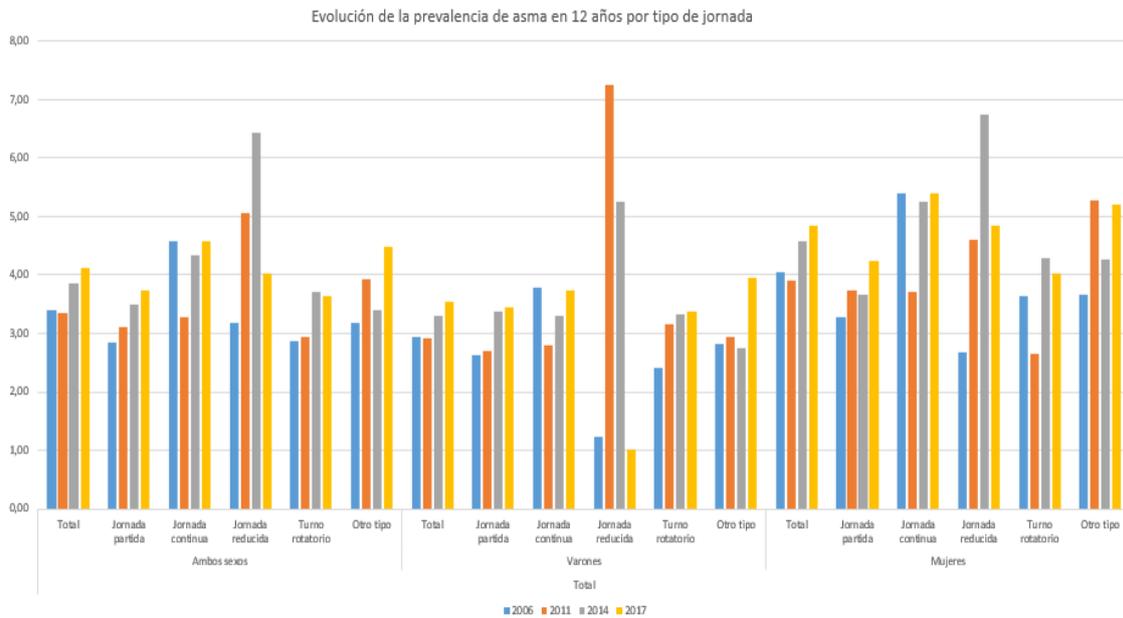


Figura 47. Diagrama de barras con el porcentaje de prevalencia en ordenadas y en abscisas los diferentes tipos de jornada con 3 divisiones: cifras totales, varones y mujeres, y a su vez cada una de ellas con subdivisiones de colores distintos según el año en que se realice la encuesta (relación de año con su color en la leyenda del pie de la figura).

La tabla 37 muestra los valores de estadística descriptiva de la prevalencia de asma en relación con el tipo de jornada. A su vez, la figura 48 grafica los valores de la media en cada uno de los nichos de la estadística descriptiva. No aparecen variaciones que sean de importancia si comparamos la evolución de la prevalencia de asma respecto al tipo de jornada laboral de los sujetos. El único dato de interés es que el punto más alto lo da el grupo de mujeres en jornada continua.

Resultados

Las desviaciones estándar están entre el 5-40% de los valores de la media. El apuntamiento de la curtosis está desplazado a la izquierda, mientras que la asimetría se reparten a ambos lados según el grupo analizado.

Tabla 37. Estadística descriptiva de la evolución de la prevalencia de asma con respecto a jornada laboral incluyendo divisiones de cifras totales, varones y mujeres.

	Total																	
	Ambos sexos					Varones					Mujeres							
	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
Media	3,687219	3,289336	4,196662	4,670324	3,295115	3,741329	3,178423	3,041724	3,402825	3,123143	3,063588	3,114653	4,340202	3,726828	4,930793	4,716186	3,653303	4,603596
Error típico	0,187394	0,198954	0,309275	0,700208	0,222727	0,291971	0,15353	0,213863	0,228967	1,848625	0,222977	0,280129	0,217442	0,195824	0,408287	0,831805	0,360818	0,389158
Mediana	3,62909	3,291073	4,455826	4,542322	3,295029	3,658008	3,123792	3,046086	3,513343	2,624034	3,242654	2,876921	4,312642	3,700908	5,314747	4,717992	3,832135	4,738522
Desviación estándar	0,374787	0,397909	0,618549	1,400417	0,445454	0,583941	0,30706	0,427727	0,457934	3,697249	0,445953	0,560258	0,434884	0,391648	0,816574	1,66361	0,721636	0,778315
Varianza de la muestra	0,140466	0,158331	0,382603	1,961167	0,198429	0,340987	0,094286	0,18295	0,209704	13,66965	0,198874	0,31389	0,189124	0,153388	0,666794	2,767597	0,520758	0,605774
Curtosis	-3,647703	-2,44023	3,115655	-0,827724	-5,779517	-1,490039	-2,997012	-5,786923	-0,963722	-4,595253	3,167121	3,641487	-3,830059	1,330214	3,869633	1,400661	1,228159	-3,165524
Coefficiente de asimetría	0,405058	-0,018038	-1,774591	0,44064	0,000199	0,602757	0,521549	-0,010213	-0,898471	0,246301	-1,782233	1,898036	0,196848	0,387878	-1,963534	-0,006473	-1,210736	-0,492364
Rango	0,76896	0,89565	1,299766	3,247341	0,835078	1,309737	0,63962	0,800236	0,985191	7,244502	0,953794	1,190642	0,926319	0,952348	1,674199	4,060215	1,645135	1,613926
Mínimo	3,360869	2,839774	3,287615	3,174655	2,877662	3,169782	2,913245	2,637244	2,799712	0	2,407626	2,757063	3,904602	3,276574	3,709739	2,680272	2,651903	3,661707
Máximo	4,129829	3,735425	4,587381	6,421996	3,71274	4,479519	3,552864	3,43748	3,784904	7,244502	3,36142	3,947705	4,83092	4,228922	5,383938	6,748486	4,297038	5,275633
Suma	14,74888	13,15734	16,78665	18,68129	13,18046	14,96532	12,71369	12,16689	13,6113	12,49257	12,25435	12,45861	17,36081	14,90731	19,72317	18,86474	14,61321	18,41438



Figura 48. Promedio de la prevalencia del asma con respecto a jornada laboral incluyendo divisiones de cifras totales, varones y mujeres.

Resultados

La figura 48 al graficar las cifras de media de la estadística descriptiva revela un pequeño aumento de la prevalencia de asma en jornada continua en varones y mujeres por separado.

La tabla 38 nos da la clave numérica para identificar cada una de las variables que aparecen en la figura 49 y que corresponden a diferentes nichos de jornada, edad y sexo.

La figura 49 es un biplot de evolución de prevalencia del asma. Cada uno de los ejes corresponde a los años de medición, situándose sobre los ejes los porcentajes de prevalencia, y sobre ellos cada uno de los grupos de jornada laboral, edad y sexo. Ocupando las posiciones más extremas están el grupo de turno rotatorio 23, 29, y 35; o con jornadas de otro tipo como 30 y 36. En el lado contrario los nichos de edades de 15-24 años como el 26 representante de la jornada partida; también separados 21, 27 y 33 que son de jornada continua.

Tabla 38. Clave de clasificación de la población (cifras totales, varones, mujeres) añadiendo los grupos de edad y el tipo de jornada laboral y numerando cada una de estas divisiones, para después ordenar sus cifras de prevalencia correspondientes.

1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18	
Total																																			
Ambos sexos						Varones						Mujeres																							
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36																		
16-24 años																																			
Ambos sexos						Varones						Mujeres																							
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54																		
25-44 años																																			
Ambos sexos						Varones						Mujeres																							
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72																		
45-64 años																																			
Ambos sexos						Varones						Mujeres																							
Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo	Total	Jornada partida	Jornada continua	Jornada reducida	Turno rotatorio	Otro tipo

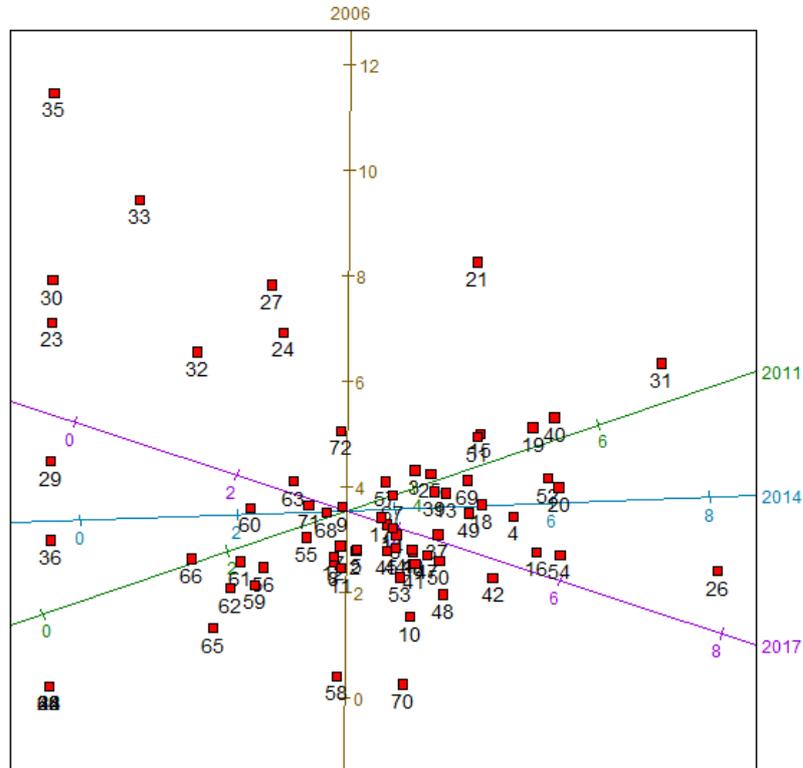


Figura 49. Biplot de representación de evolución de asma según tipo de jornada laboral.

4.2. Conglomerados o *clusters*

4.2.1. Análisis de conglomerados de la evolución de la prevalencia de alergia en 25 años

◆ CONGLOMERADO DE LA PREVALENCIA DE ALERGIA POR EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en *clusters* y su representación gráfica. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos.

Clustering vector: 4 2 2 4 4 3 3 1 3 2 4 4 1 1 1 1 2 2 2 2 2 2 4 3

Resultados

Tabla 39. Clave numérica que muestra ordenados los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.

1	2	3	4	5	6	7	8
Ambos sexos							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años
10,003205	12,7179252	11,3492223	9,97162886	9,27984452	9,08876394	8,35003941	7,30123361
4	2	2	4	4	3	3	1
9	10	11	12	13	14	15	16
Varones							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años
8,5257	12,5140329	10,2970914	8,53656218	6,96471141	6,67878844	5,82412608	5,31686154
3	2	4	4	1	1	1	1
17	18	19	20	21	22	23	24
Mujeres							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años
11,4044685	12,9420273	12,4161483	11,4247209	11,5341343	11,3565582	10,3495367	8,71713965
2	2	2	2	2	2	4	3

En la tabla 39 podemos observar en que conglomerado quedan agrupados cada uno de los nichos de edad y sexo con sus respectivas cifras de prevalencia de alergia.

El conglomerado que solo tiene 4 componentes que es el número 3 comprende los nichos de: totales de varones y los totales de ambos sexos de 55-64 años y de 65-74 años, así como, el de las mujeres de 75 años y más. Serían las prevalencias medias.

El *cluster* 4 que tiene 6 componentes también agruparía estas prevalencias medias. El conglomerado que tiene 5 componentes, que es el número 1 comprende los totales de más de 75 años junto con los varones de 45-54 años, 55-64 años, 65-74 años y más de 75 años. Correspondería a las prevalencias más bajas.

El más numeroso que es el conglomerado 2 comprende toda la serie de mujeres excepto las de más edad de 65-74 años y de 75 y más años. Serían las prevalencias más altas. Es decir, los conglomerados está seleccionando por mayor prevalencia y por menor prevalencia.

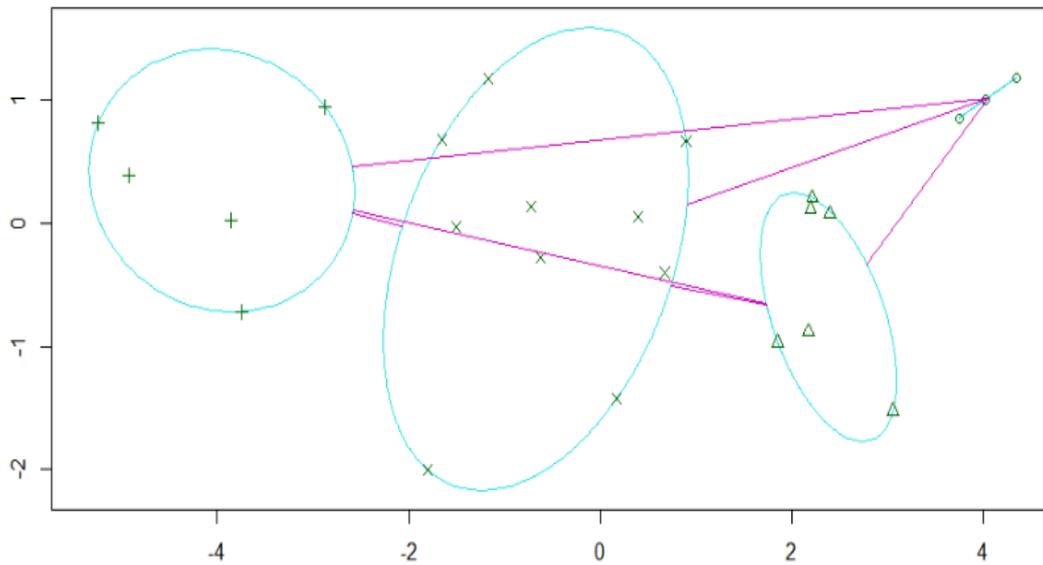


Figura 50. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia en 25 años por edad y sexo.

4.2.2. Análisis de conglomerados de la evolución de la prevalencia de alergia según tipo de población

◆ CONGLOMERADO DE LA PREVALENCIA DE ALERGIA POR POBLACIÓN, EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en conglomerados y su representación gráfica. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos. Se obtienen 4 conglomerados de tamaños 35, 13, 25, 17.

Clustering vector: 1 1 1 1 1 2 3 3 3 3 1 1 1 1 1 2 4 2 4 4 4 4 4 2 4 4 4 4 2 4 4
 4 2 4 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 4 2 1 3 1 3 1 1 3 3 3 3 3 1 1 1 2 4 2 3 3
 3 3 3 3 3 3 3 3 3 1 3 1 2 1 2

La figura 51 y la tabla 40 nos muestra la agrupación que presentan cada uno de los grupos de edad, sexo y tipo de población.

El conglomerado 2 agrupa únicamente a los alérgicos en ciudades de más de 400.000 habitantes sin que sean semejantes las cifras de prevalencia.

El conglomerado número 4 concentra a la población alérgica de 15-24 años independientemente del sexo exceptuando las poblaciones de más de 400.000 habitantes, y agruparía grupos de alta prevalencia.

El conglomerado 1 agrupa a los nichos de 25-44 años y algunos de 45-64 años y el conglomerado 3 a los grupos de mayor edad.

Es decir, los conglomerados nos servirían para diferenciar grupos de distinta prevalencia e incluso agruparía a poblaciones con mayor importancia al respecto.

La estadística descriptiva nos habla de que por poblaciones vemos que en los últimos años ha aumentado sobre todo la prevalencia en las de 100.000 a 400.000 habitantes y en poblaciones de más de 400.000 habitantes. Mientras que en el resto de poblaciones más pequeñas las cifras son más similares y con menos diferencias entre ellas. Esto es más acusado en la población femenina.

Por tanto, el *cluster* 2 nos estaría agrupando esta tendencia ya que agrupa únicamente a los alérgicos en ciudades de más de 400.000 habitantes sin que sean semejantes las cifras de prevalencia.

Resultados

Tabla 40. Clave que muestra ordenados numéricamente los diferentes nichos de población, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.

Total																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	1	1	1	1	2	3	3	3	3	1	1	1	1	1	1	2	4	2
10,00361	8,90095	9,520487	10,0779	10,49123	11,52358	8,526253	7,946747	8,111398	8,333802	8,804274	9,700698	11,40466	9,823792	10,87499	11,76493	12,06161	13,20653	
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
15-24 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
4	4	4	4	4	2	4	4	4	4	4	2	4	4	4	4	2	4	2
12,71745	11,47331	12,30807	13,13154	13,50114	13,77768	12,51401	11,36681	12,52862	12,78668	13,02376	13,26197	12,94182	11,46643	12,17652	13,94048	13,98842	14,46581	
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
25-44 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	4	2
10,66662	9,834052	9,809565	10,37104	11,18179	12,32409	9,48078	9,127263	8,489598	9,212488	9,863183	10,95038	11,87563	10,56602	11,15912	11,60489	12,50726	13,65916	
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
45-64 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	3	1	3	1	3	1	1	3	3	3	3	3	3	3	1	1	1	2
9,201413	8,668378	8,721882	8,555108	9,255954	10,92917	6,8556	7,153522	6,580084	5,337623	6,887514	7,830664	11,4516	10,21844	10,80569	11,67485	11,54875	13,76498	
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	
65 y más años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
4	2	3	3	3	3	3	3	3	3	3	3	3	1	3	1	2	1	2
7,894307	6,590968	7,909648	8,747018	8,68151	8,714658	5,599818	5,201677	5,807555	6,108601	5,455632	5,835309	9,598789	7,707231	9,500364	10,36195	10,93728	10,72922	

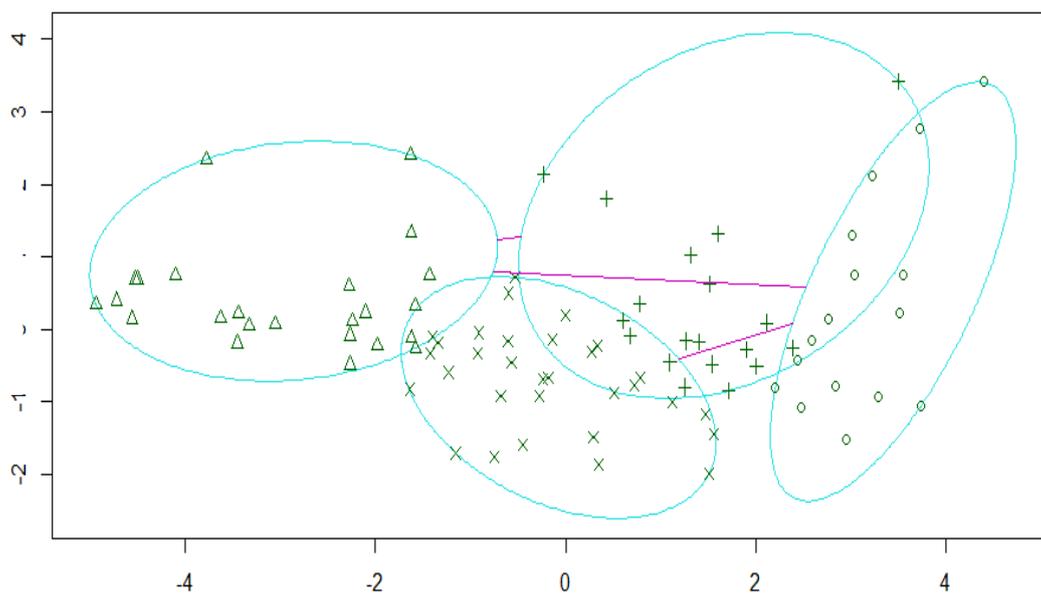


Figura 51. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia según tipo de población junto por edad y sexo.

4.2.3. Análisis de conglomerados de la prevalencia de alergia según nivel de estudios

◆ CONGLOMERADO DE LA PREVALENCIA DE ALERGIA POR ESTUDIOS, EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en *clusters* y su representación gráfica. Se obtiene 3 conglomerados con 29, 43, 18 componentes. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos.

Resultados

Clustering vector: 1 1 1 1 2 2 1 3 3 1 1 2 2 2 1 2 2 2 2 1 2 2 2 2 3 2 1 2 2 2 2
2 2 2 2 2 1 2 2 2 1 3 3 1 2 2 2 2 2 2 2 1 1 1 1 1 2 3 3 3 1 1 2 2 2 2 1 2 3 1
3 1 3 1 3 3 3 3 3 1 1 1 1 1 2

La tabla 41 y la figura 52 nos muestra la agrupación que presentan cada uno de los grupos de edad, sexo y nivel de estudios.

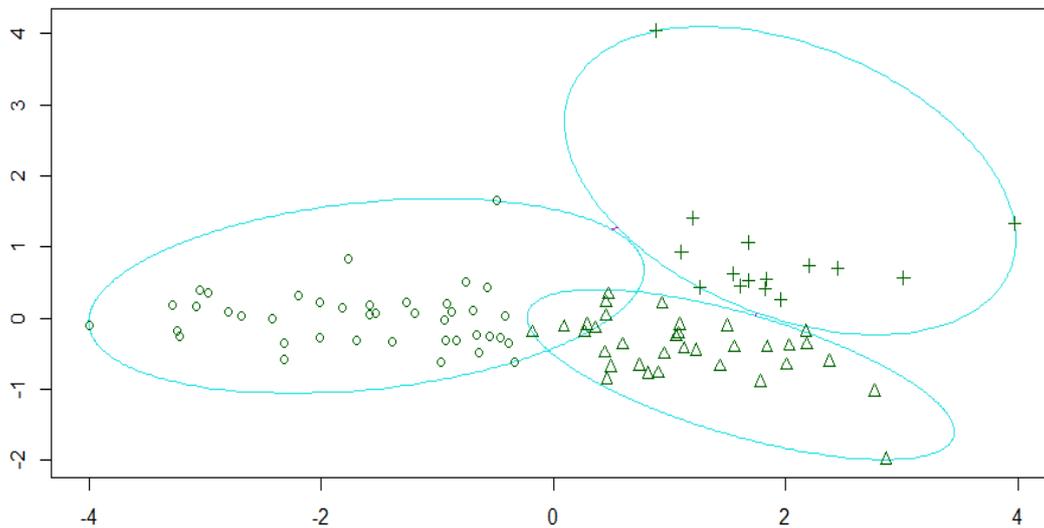


Figura 52. Representación gráfica de los conglomerados en que queda dividida la distribución de alergia en niveles de estudios y por edad y sexo.

El conglomerado 2 incluye a las mujeres de grupos de edad de 15-24 años y de 25-44 años y algunos nichos de 45-64 años. También a los varones más jóvenes y con mayor nivel educativo.

El conglomerado 3 incluye a los nichos de varones sin estudios y a las poblaciones de mayor edad, de 45-64 años y más de 65 años (sobre todo varones), siendo todos ellos los que muestran las prevalencias más bajas. El conglomerado 1 incluye a estas poblaciones pero sobre todo mujeres.

Resultados

La estadística descriptiva nos presenta que los valores de alergia son más altos en personas con niveles superiores de estudios tanto en cifras totales como si separamos en hombres y mujeres, ocurriendo este hecho independientemente del año en que se analice. El conglomerado 2 nos recogería a estos nichos.

Tabla 41. Clave que muestra ordenados numéricamente los diferentes nichos de niveles de estudios, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.

Total																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	1	1	1	2	2	1	3	3	1	1	2	2	2	2	1	2	2	2
12,480722	10,842268	10,876699	12,083598	13,300934	14,71079	10,764615	7,9018391	8,6991321	10,103682	11,723335	13,704793	14,117386	13,136647	12,787447	14,180035	14,861686	15,681398	
15-24 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
2	1	2	2	2	2	2	3	2	1	2	2	2	2	2	2	2	2	2
14,967956	11,225312	14,845424	13,821202	15,24287	17,111523	14,702905	8,3730536	14,236598	12,984748	14,336784	18,436628	15,246456	13,306028	15,528055	14,793051	16,200806	15,456461	
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
25-44 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
2	2	1	2	2	2	1	3	3	1	2	2	2	2	2	2	2	2	2
14,214779	13,336039	11,483301	13,148925	15,298225	15,815016	12,741923	9,3204742	9,7373286	11,337272	13,715967	15,342321	15,728275	17,470984	13,227624	15,061249	16,903343	16,250692	
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
45-64 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	1	1	1	1	2	3	3	3	3	1	1	2	2	2	2	1	2	2
11,612655	11,084953	11,006942	10,860472	11,259481	13,873382	9,0981863	7,8217221	7,1656286	8,4108175	9,6287518	11,761618	14,063173	14,036832	14,309006	13,614836	12,897106	15,876041	
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	
65 y más años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
3	1	3	1	3	1	3	3	3	3	3	3	3	1	1	1	1	1	2
9,3390644	9,7478378	8,9832968	9,7668634	8,4912871	10,235824	6,9241504	7,0809167	7,2267998	7,437478	5,6616309	7,3669933	11,186174	11,405127	10,35961	11,6004	10,776981	13,123912	

4.2.4. Análisis de conglomerados de la evolución de la prevalencia de asma

◆ CONGLOMERADO DE LA PREVALENCIA DE ASMA POR EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en *clusters* y su representación gráfica. Se obtiene 3 conglomerados de tamaño 15, 3, 6. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos.

Clustering vector: 1 1 1 1 2 1 1 3 1 1 1 1 2 2 1 3 1 3 3 1 1 1 3 3

La figura 53 y la tabla 42 nos muestra la agrupación que presentan cada uno de los grupos de edad y sexo.

Tabla 42. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico, con sus respectivas numeraciones de pertenencia a un determinado *cluster* (localizadas en la zona de verde más intenso). Así mismo, aparecen las cifras de prevalencia.

	1	2	3	4	5	6	7	8
Ambos sexos								
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	1	1	1	1	2	1	1	3
	4,392593	4,729988	4,766326	3,977654	3,330029	3,8534	4,96207	6,244296
Varones								
	9	10	11	12	13	14	15	16
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	1	1	1	1	2	2	1	3
	3,736262	4,162501	4,204529	3,669307	2,410279	2,868817	4,156797	6,245398
Mujeres								
	17	18	19	20	21	22	23	24
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	1	3	3	1	1	1	3	3
	5,018262	5,323815	5,346312	4,297392	4,246525	4,789411	5,656427	6,226391

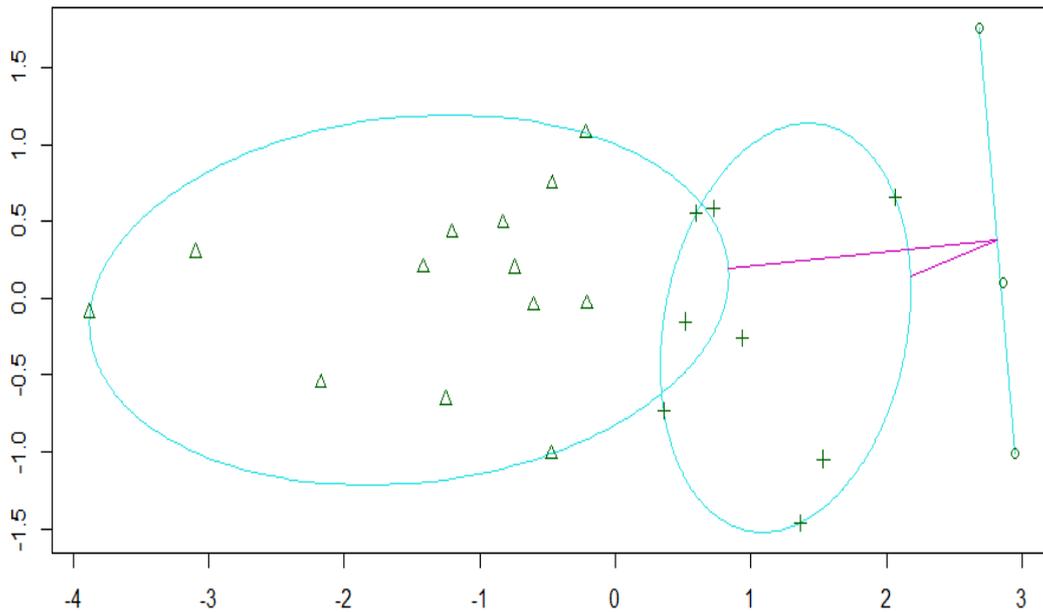


Figura 53. Representación gráfica de los conglomerados en que queda dividida la distribución del asma por edad y sexo.

El conglomerado 3 integra a los nichos de mayor edad y a las mujeres de 15-24 y de 25-34 años, que son precisamente los que presentan una mayor prevalencia.

El conglomerado 2 incluye a los varones de 45-54 años y a los 55-64 años que son los de menor prevalencia, junto con el grupo de 45-54 de ambos sexos que dentro de este apartado global es el de menor prevalencia.

4.2.5. Análisis de conglomerados de la evolución de la prevalencia de asma según tipo de población

- ◆ CONGLOMERADO DE LA PREVALENCIA DE ASMA POR POBLACION, EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en *clusters* y su representación gráfica. Se obtienen 3 conglomerados de tamaños 22,

45, 23. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos.

Clustering vector: 2 2 2 2 2 2 1 2 2 1 1 1 2 2 2 2 3 2 2 3 2 1 3 3 2 2 2 1 2 1 3 3
 2 1 3 3 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 3 2 1 1 1 1 2 1 1 1 1 1 1 1 2 1 2 3 2 2 3 3
 3 3 3 2 3 3 3 2 2 1 3 3 3 3 2

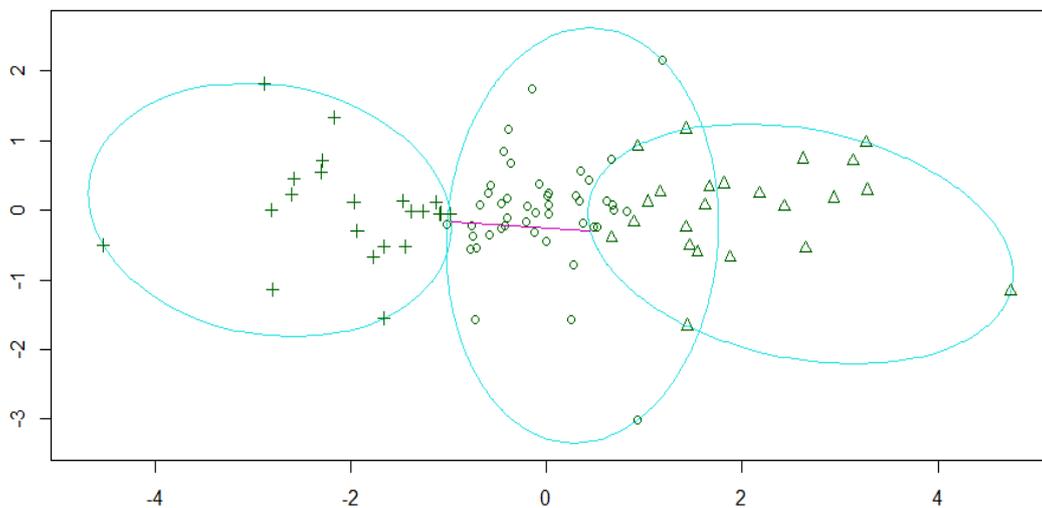


Figura 54. Representación gráfica de los *clusters* en que queda dividida la distribución del asma según tipo de población junto por edad y sexo.

La figura 54 y las tabla 43 nos muestran como quedan integrados cada uno de los nichos de edad, sexo y tipo de población en los diferentes conglomerados. Además, la tabla 43 incorpora las cifras de prevalencia para facilitar la comparativa.

El conglomerado 3 integra a las poblaciones de mayor edad de más 65 años y varios nichos de la población más joven de 15-24 años independientemente del tipo de población. El *cluster* 2 une a los grupos más jóvenes (15-24 y 25-44 años) independientemente de la población. En conclusión, estos datos de población no son relevantes en la composición de los conglomerados.

Resultados

La estadística descriptiva nos muestra que en poblaciones de más de 400.000 habitantes se dan las cifras más bajas en totales de ambos sexos y también en mujeres. Las cifras más altas son en poblaciones de 100.000 a 400.000 habitantes aunque en el caso de varones las cifras más altas corresponden a poblaciones de hasta 10.000 habitantes.

Tabla 43. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado. Así mismo, aparecen las cifras de prevalencia.

Total																	
Ambos sexos						Varones					Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
4,3925934	4,4899221	4,461459	4,1798222	4,6854059	3,9500351	3,7362623	4,0864097	3,9877424	3,3803552	3,7196193	3,1961606	5,0182625	4,8914829	4,9094715	4,9491508	5,5926942	4,638938
2	2	2	2	2	2	1	2	2	1	1	1	1	2	2	2	2	3
15-24 años																	
Ambos sexos						Varones					Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
4,7299883	5,1172289	4,2377186	2,9569726	5,6643706	5,2747441	4,1625007	5,2032251	4,2040156	3,2985915	4,6318051	3,0917275	5,3238151	5,020879	4,2338129	2,5773505	6,7029493	7,8299181
2	2	3	2	1	3	3	2	2	2	1	2	1	3	3	2	1	3
25-44 años																	
Ambos sexos						Varones					Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
4,3480452	4,3204537	4,2868269	4,4154853	4,7683173	3,8818199	3,9275609	4,1860638	3,9134929	4,0879633	4,106064	3,3280591	4,7841335	4,4589325	4,670302	4,750575	5,4591448	4,4763782
3	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2
45-64 años																	
Ambos sexos						Varones					Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
3,5587592	3,178876	3,6909591	3,5788626	3,7625466	3,5361869	2,6090656	2,7542657	2,9308643	1,5311983	2,5759612	2,8052713	4,4856558	3,6466342	4,4259802	5,5589482	4,8829539	4,2158895
3	2	1	1	1	1	1	1	1	1	1	1	2	1	2	3	2	2
65 y más años																	
Ambos sexos						Varones					Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
5,5784791	6,124168	6,4529432	5,4044955	5,4633022	3,9917977	5,1000749	5,5015323	6,2729335	5,0214804	4,4094434	3,8522919	5,9456017	6,6497072	6,5751186	5,7333283	6,2662539	4,1117431
3	3	3	3	3	2	3	3	3	2	2	1	3	3	3	3	3	3

4.2.6. Análisis de conglomerados de la evolución de la prevalencia de asma según nivel de estudios

◆ CONGLOMERADO DE PREVALENCIA DE ASMA POR ESTUDIOS, EDAD Y SEXO

Utilizando los algoritmos k-means y clusplot de R obtenemos una agrupación en conglomerados y su representación gráfica. Se obtienen 3 conglomerados de tamaños 53, 16, 21. Las variables quedan adjudicadas, siguiendo el orden de la clave numérica, en cada uno de los conglomerados a continuación referidos.

Clustering vector: 1 2 1 1 1 1 1 2 1 1 3 1 1 2 1 1 1 1 1 3 2 2 1 1 1 3 1 1 3 3 2 3
 2 2 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 2 1 1 2 1 1 3 3 3 3 3 3 3 1 2 1 1 1 3 2 2
 1 1 3 3 1 2 1 3 3 3 2 2 1 1 1 1

La figura 55 y la tabla 44 nos muestra la agrupación que presentan cada uno de los grupos de edad y sexo en el caso de asma y nivel de estudios. Además, la tabla 44 incorpora las cifras de prevalencia para facilitar la comparativa.

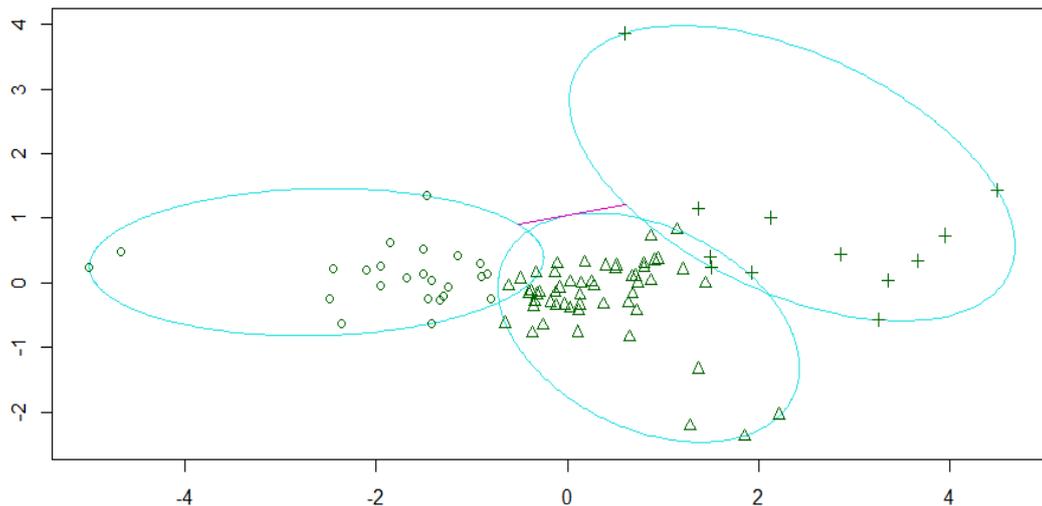


Figura 55. Representación gráfica de los conglomerados en que queda dividida la distribución del asma por nivel de estudios junto por edad y sexo.

El grupo de 25-44 años es el más homogéneo de todos con pertenencia casi exclusiva al *cluster* 1.

En el conglomerado de menos miembros (16 en total) que es el 2 están incluidos los siguientes grupos (numerados según la clave):

2 que corresponde a total ambos sexos sin estudios, 8 total varones sin estudios, 14 total mujeres sin estudios, 21 ambos sexos de 15-24 años primer grado, 22 ambos sexos de 15-24 años segundo grado primer ciclo, 33 mujeres de 15-24 años primer grado, 56 ambos sexos de 45-64 años sin estudios, 73 ambos sexos total de 65 y más años, 79 total varones de más de 65 años, 86 mujeres de más de 65 años sin estudios.

Es decir, que el conglomerado 2 incluiría varios grupos que coincidirían con lo reflejado en la estadística descriptiva. Hay un marcado aumento en la población sin estudios, tanto en ambos sexos como por separado. El grupo con estudios superiores tendría las prevalencias menores.

Resultados

Tabla 44. Clave que muestra ordenados numéricamente los diferentes nichos de niveles de estudios, edad y sexo que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de verde más intenso). Así mismo, aparecen las cifras de prevalencia.

Total																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	2	1	1	1	1	1	2	1	1	3	1	1	2	1	1	1	1	1
4,39259	6,76674	4,5008	4,19257	3,92704	3,84648	3,73626	6,16987	4,26544	3,56513	2,99407	3,49986	5,01826	7,17783	4,71406	4,88752	4,90138	4,15292	
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
15-24 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	3	2	2	1	1	1	3	1	1	3	3	2	3	2	2	1	1	
4,72999	0,3082	5,59048	5,10079	4,00162	5,21852	4,1625	0	6,01283	4,67433	2,99611	2,1958	5,32382	0	5,09384	5,53994	4,97417	5,72355	
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
25-44 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	1	1	1	1	1	1	3	1	1	1	1	1	1	1	1	2	2	
4,34805	4,01476	4,11842	4,28144	4,56121	4,29545	3,92756	2,99524	4,25975	3,87576	3,65124	4,19154	4,78413	5,39975	3,95445	4,73287	5,51196	4,37475	
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
45-64 años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
1	2	1	3	3	3	3	3	3	3	3	3	1	2	1	1	1	3	
3,55876	5,35802	4,00728	3,51253	3,04148	2,99782	2,60907	3,13861	3,1589	2,58362	2,1741	2,69733	4,48566	7,06152	4,81578	4,46964	3,97019	3,30908	
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	
65 y más años																		
Ambos sexos						Varones						Mujeres						
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	
2	1	1	3	3	1	2	1	1	3	3	3	2	2	1	1	1	1	
5,57848	7,71759	4,96162	3,91191	3,12074	3,18243	5,10007	8,22479	4,83298	2,85462	2,16612	2,48357	5,9456	7,45481	5,06413	4,83967	4,22025	4,29545	

4.3. ANÁLISIS DE COMPONENTES PRINCIPALES

El análisis de componentes principales es una de las técnicas de aprendizaje no supervisado, las cuales suelen aplicarse como parte del análisis exploratorio de los datos. A diferencia de los métodos de aprendizaje supervisado, donde contamos con un grupo de variables medidas sobre un conjunto de observaciones con la intención de obtener predicciones sobre una variable respuesta, en los métodos no supervisados solo contamos con un grupo de variables de las cuales nos interesa conocer o de las que queremos extraer información, por ejemplo sobre la existencia de subgrupos entre las variables u observaciones.

Una de sus aplicaciones es la reducción de la dimensionalidad (variables), perdiendo la menor cantidad de información (varianza) posible: cuando contamos con un gran número de variables cuantitativas posiblemente correlacionadas (indicativo de información redundante), el análisis de componentes principales permite reducir las a un número menor de variables transformadas (componentes principales que expliquen gran parte de la variabilidad de los datos). Se utiliza para sacar a relucir patrones fuertes de un conjunto de datos y para que los datos sean fáciles de visualizar y explorar.

Los autovectores son los coeficientes de la ecuación de cada componente principal. Los autovalores son los cuadrados de las desviaciones estándar. La proporción de varianza es lo que explica a cada componente principal, siendo siempre la suma igual al valor uno. Los componentes principales son proyecciones que muestran donde los datos están más extendidos y presentan una mayor varianza. Un biplot es una generalización del diagrama de dispersión simple de dos variables y permite mostrar gráficamente información sobre muestras y variables de una matriz de datos. El círculo de correlaciones es útil para interpretar el significado de los ejes. Los vectores de cada parámetro se presentan en la gráfica frente a los dos componentes principales rotados.

4.3.1. Análisis de componentes principales de la evolución de la prevalencia de alergia en 25 años

- ANÁLISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ALERGIA POR EDAD Y SEXO

Tabla 45. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico.

	1	2	3	4	5	6	7	8
	Ambos sexos							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	9	10	11	12	13	14	15	16
	Varones							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	
	17	18	19	20	21	22	23	24
	Mujeres							
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años	

Los componentes principales son proyecciones que muestran donde los datos están más extendidos y presentan una mayor varianza. El círculo de correlaciones es útil para interpretar el significado de los ejes. Cuanto más se acerca la flecha de la variable al perímetro del círculo mayor es la significatividad de la variable.

Vemos en la figura 56 que las variables se agrupan en el eje horizontal, y en la zona positiva del componente 2, no mostrando ninguna variable correlación negativa. Por tanto, son variables en que los vectores no muestran resultados oponentes. La longitud de los vectores es similar por lo que las variables muestran poca diferencia entre ellas. Se correlacionan positivamente porque existe poco ángulo entre ellas. Los grupos de edad entre 35-44 años se agrupan en el cuarto superior independientemente del sexo.

Los grupos 2, 10 y 18 que corresponderían al nicho de edad de 15-24 años y por tanto los de mayor prevalencia se sitúan muy próximos en el cuadrante inferior.

Resultados

Los grupos de edad más alta como el de 75 años (clave 16 y 24) ocupan los extremos superior e inferior de los cuartos, estando con un ángulo que no llega a ser ortogonal pero es el que más se acerca de toda la distribución y esto nos hablaría de una escasa relación. El 16 es el grupo de 75 años de varones y el 24 el de mujeres y ambos son los de menor prevalencia entre su grupo de género, hecho que también es indicado por una menor longitud de la flecha de la variable.

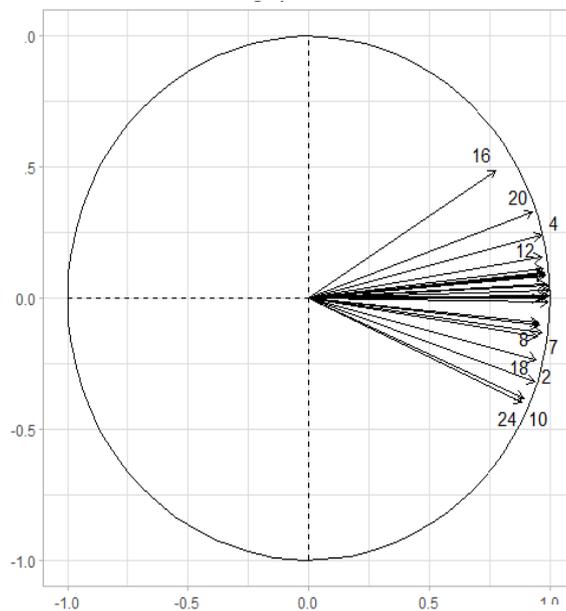


Figura 56. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave de la tabla 45.

Resultados

4.3.2. Análisis de componentes principales de la evolución de la prevalencia de alergia en 25 años según el tipo de población

- ANALISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ALERGIA POR POBLACION, EDAD Y SEXO

La tabla 46 nos transmite la clave para interpretar la representación gráfica del círculo de correlaciones y muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con tipo de población que son utilizados para el análisis estadístico.

Tabla 46. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con tipo de población que son utilizados para el análisis estadístico con sus respectivas numeraciones de pertenencia a un determinado conglomerado (localizadas en la zona de azul más intenso). Así mismo, aparecen las cifras de prevalencia.

Total																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	1	1	1	1	2	3	3	3	3	3	1	1	1	1	1	2	4	2
10,00361	8,90095	9,520487	10,0779	10,49123	11,52358	8,526253	7,946747	8,111398	8,333802	8,804274	9,700698	11,40466	9,823792	10,87499	11,76493	12,06161	13,20653	
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
15-24 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
4	4	4	4	4	2	4	4	4	4	4	4	2	4	4	4	2	4	2
12,71745	11,47331	12,30807	13,13154	13,50114	13,77768	12,51401	11,36681	12,52862	12,78668	13,02376	13,26197	12,94182	11,46643	12,17652	13,94048	13,98842	14,46581	
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
25-44 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	4	2	
10,66662	9,834052	9,809565	10,37104	11,18179	12,32409	9,48078	9,127263	8,489598	9,212488	9,863183	10,95038	11,87563	10,56602	11,15912	11,60489	12,50726	13,65916	
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
45-64 años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
1	3	1	3	1	3	1	1	3	3	3	3	3	3	3	1	1	1	2
9,201413	8,668378	8,721882	8,555108	9,255954	10,92917	6,8556	7,153522	6,580084	5,337623	6,887514	7,830664	11,4516	10,21844	10,80569	11,67485	11,54875	13,76498	
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	
65 y más años																		
Ambos sexos						Varones						Mujeres						
Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	Total	Hasta 10.000 habitantes	De 10.001 a 50.000 habitantes	De 50.001 a 100.000 habitantes	De 100.001 a 400.000 habitantes	Más de 400.000 habitantes	
4	2	3	3	3	3	3	3	3	3	3	3	3	1	3	1	2	1	2
7,894307	6,590968	7,909648	8,747018	8,68151	8,714658	5,599818	5,201677	5,807555	6,108601	5,455632	5,835309	9,598789	7,707231	9,500364	10,36195	10,93728	10,72922	

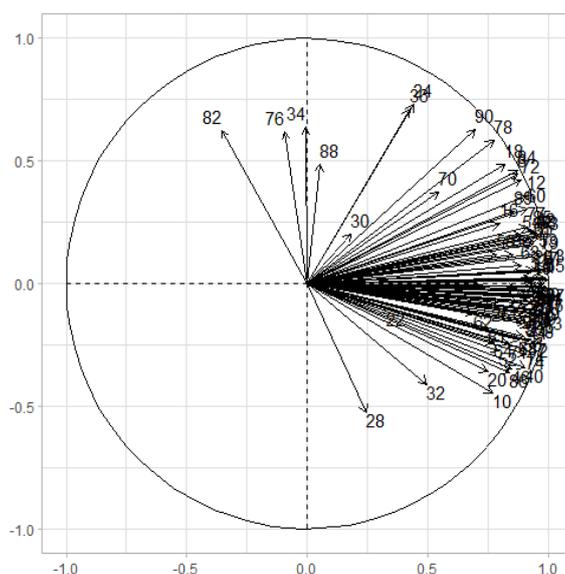


Figura 57. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.

El círculo de correlaciones de la figura 57 muestra a las poblaciones en que la prevalencia es mayor, según la estadística descriptiva, acumuladas en la zona positiva del componente 1.

Las variables se correlacionan positivamente porque existe poco ángulo entre ellas. La longitud de los vectores en su mayoría es similar por lo que las variables muestran poca diferencia entre ellas. Las variables se agrupan en el eje horizontal del principal componente, no mostrando ninguna variable correlación opuesta entre ellas, excepto la 28 (varones de 15-24 años en poblaciones de 5.0000-100000 habitantes) con la 82 (varones del mismo tipo de población de más de 65 años), presentando diferencias de prevalencia entre ellos: 12,78% frente 6,10%.

Alrededor del eje vertical se agrupan las variables de poblaciones entre 50.000-100.000 habitantes: 82 (6,10%) varones de más de 65 años, 76 (8,74%) totales de ambos sexos de más de 65 años, 34 (13,94%) mujeres de 15-24 años, todas ellas de poblaciones de 50.000-100.000 habitantes. Como se aprecia no hay semejanza en las prevalencias entre ellas.

A su vez, el 32 (11,46%) es el grupo de mujeres valorado de 15-24 años en poblaciones de menos de 10.000 habitantes que expresaría correlación negativa al estar en el eje vertical inferior.

Dado que cuanto más se acerca la flecha de la variable al perímetro del círculo mayor es la significatividad de la variable, una menor longitud de la flecha de la variable indica menor prevalencia y menor significatividad de estas variables. Esto coincidiría con la estadística descriptiva que nos comunica la menor relevancia de estos grupos. La estadística descriptiva nos habla de que por poblaciones vemos que en los últimos años ha aumentado sobre todo la prevalencia en poblaciones de 100.000 a 400.000 habitantes y en poblaciones de más de 400.000 habitantes. Mientras que en el resto de poblaciones más pequeñas las cifras son más similares y con menos diferencias entre ellas. Esto es más acusado en la población femenina.

4.3.3. Análisis de componentes principales de la evolución de la prevalencia de alergia según nivel de estudios

- ANÁLISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ALERGIA POR ESTUDIOS, EDAD Y SEXO

La tabla 47 nos transmite la clave para interpretar la representación gráfica del círculo de correlaciones de la figura 58 y muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el nivel de estudios que son utilizados para el análisis estadístico.

El círculo de correlaciones de la figura 58 muestra situación de los diferentes nichos de la tabla 47. La longitud de los vectores es similar por lo que las variables muestran poca diferencia entre ellas.

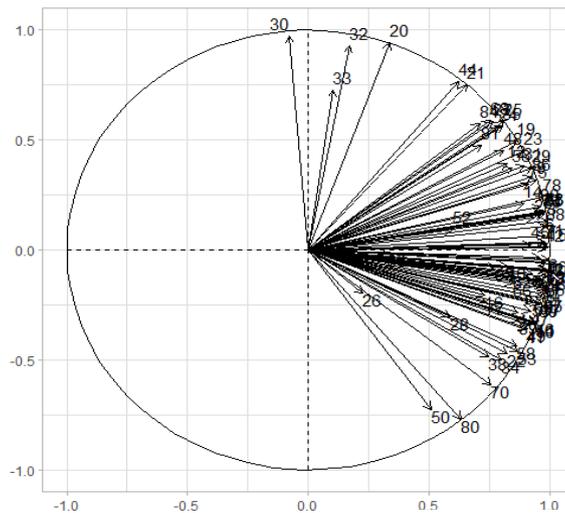


Figura 58. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.

Las variables se agrupan en los dos cuadrantes de la zona del componente principal que explica la mayor parte de la varianza, no mostrando ninguna variable correlación negativa excepto la clave 30 que corresponde al grupo de varones de 15-24 años con estudios de 3º (18,43%).

La mayoría de variables se confunden en su trazado al haber poco ángulo entre ellas lo cual quiere decir que se correlacionan positivamente. Únicamente las claves 20 nicho de ambos sexos sin estudios de 15-24 años (11,22%), 32 mujeres de 15-24 años sin estudios (13,30%) y 33 mujeres de 15-24 años primer grado (15,52%), muestran un ángulo recto con respecto al resto de variables expresando su nula correlación. Se trataría de nichos con una prevalencia similar respecto al resto, por lo que no es el factor que determina. Su separación nos serviría para determinar visualmente su menor significatividad en base a la menor longitud de la variable en el caso de la 33.

La falta de significatividad es evidente en la variable 26 con longitud mínima que corresponde a varones sin estudios de 15-24 años con prevalencia baja (8,37%).

Resultados

Si tenemos en cuenta los niveles de estudio considerando como tales: sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado; podemos ver que los valores de alergia son más altos en personas con nivel de tercer grado tanto en cifras totales como si separamos en hombres y mujeres, ocurriendo este hecho independientemente del año en que se analice.

Tabla 47. Clave que muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con nivel de estudios que son utilizados para el análisis estadístico. Incluye cifras de prevalencia.

Total																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
1	1	1	1	2	2	1	3	3	1	1	2	2	2	1	2	2	2
12,48072	10,84227	10,8767	12,0836	13,30093	14,71079	10,76461	7,901839	8,699132	10,10368	11,72333	13,70479	14,11739	13,13665	12,78745	14,18003	14,86169	15,6814
15-24 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
2	1	2	2	2	2	2	3	2	1	2	2	2	2	2	2	2	2
14,96796	11,22531	14,84542	13,8212	15,24287	17,11152	14,7029	8,373054	14,2366	12,98475	14,33678	18,43663	15,24646	13,30603	15,52606	14,79305	16,20081	15,45646
25-44 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
2	2	1	2	2	2	1	3	3	1	2	2	2	2	2	2	2	2
14,21478	13,33604	11,4833	13,14892	15,29823	15,81502	12,74192	9,320474	9,737329	11,33727	13,71597	15,34232	15,72827	17,47098	13,22762	15,06125	16,90334	16,25069
45-64 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
1	1	1	1	1	2	3	3	3	3	1	1	2	2	2	2	1	2
11,61265	11,08495	11,00694	10,86047	11,25948	13,87338	9,098186	7,821722	7,165629	8,410817	9,628752	11,76162	14,06317	14,03683	14,30901	13,61484	12,89711	15,87604
65 y más años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, primer ciclo	Segundo grado, segundo ciclo	Tercer grado
3	1	3	1	3	1	3	3	3	3	3	3	1	1	1	1	1	2
9,339064	9,747838	8,983297	9,766863	8,491287	10,23582	6,92415	7,080917	7,2268	7,437478	5,661631	7,366993	11,18617	11,40513	10,35961	11,6004	10,77698	13,12391

4.3.4. Análisis de componentes principales de la evolución de la prevalencia de asma

- ANÁLISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ASMA POR EDAD Y SEXO

La tabla 48 nos transmite la clave para interpretar la representación gráfica del círculo de correlaciones de la figura 59 y muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico.

Tabla 48. Muestra ordenados numéricamente los diferentes nichos de edad y sexo que son utilizados para el análisis estadístico e incorpora cifras de prevalencia.

1		2		3		4		5		6		7		8	
Ambos sexos															
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años								
4,392593	4,729988	4,766326	3,977654	3,330029	3,8534	4,96207	6,244296								
9															
10															
11															
12															
13															
14															
15															
16															
Varones															
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años								
3,736262	4,162501	4,204529	3,669307	2,410279	2,868817	4,156797	6,245398								
17															
18															
19															
20															
21															
22															
23															
24															
Mujeres															
Total	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75 y más años								
5,018262	5,323815	5,346312	4,297392	4,246525	4,789411	5,656427	6,226391								

En el círculo de correlaciones de la figura 59 es evidente la diferencia de distribución que veíamos con respecto a la alergia. La longitud de los vectores no es similar por lo que las variables muestran diferencia entre ellas. Las variables están poco correlacionadas porque existe un mayor ángulo entre ellas que en el caso de la alergia. Se distribuyen repartidas entre los dos componentes principales que explican la mayor parte de la varianza, con un pequeño aumento en el primer componente principal con respecto al segundo. Los nichos 4, 12, 20 representan el grupo de edad de 35-44 años en todas sus versiones (ambos sexos, mujeres,

varones) situándose todos ellos en la zona negativa del componente principal. En el lado contrario los numerados como 8, 16 y 24 con longitud importante y siendo todos los de edad de más de 75 años. Es decir, los menos prevalentes frente a los más prevalentes. Con proyección ortogonal tenemos al nicho 12 varones y 20 mujeres del mismo grupo de edad 35-44 años, es decir con nula correlación entre ellas.

En el eje horizontal tenemos la clave 13 (varones de 45-54 años) y el nicho 8 (total en ambos sexos de más de 75 años) con correlación opuesta, es decir con una prevalencia que difiere entre ellas, con el 4 que es total ambos sexos de 35-44 años. En el cuadrante superior derecho la clave 19 (mujeres de 25-34 años), el nicho 21 (que incluye a mujeres 45-54 años) y el 18 (mujeres de 15-24 años); todos ellos en proyección ortogonal con 14, 15 y 16, que son nichos de varones de 55-64, 65-74 y más de 75 años.

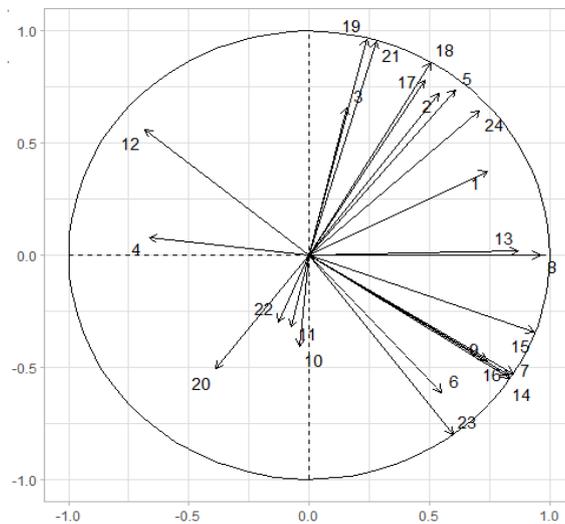


Figura 59. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.

El aumento en la población de más de 75 años tanto en cifras totales como particularmente en varones y mujeres es evidente así como la caída en grupos de 45-54 años (serían los grupos 5, 13, 21). La forma de la curva descrita en la estadística descriptiva se mantiene en las series temporales sin cambios según el

año estudiado. En todos los años medidos, en los dos sexos hay un incremento en los grupos más jóvenes con respecto a edades medias de 45-54 años, para luego verse superado el resultado claramente por encima de los 65 años. Es decir, no hay una subida constante y prolongada desde las primeras edades de la vida.

4.3.5. Análisis de componentes principales de la evolución de la prevalencia de asma según el tipo de población

- ANÁLISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ASMA POR POBLACIÓN, EDAD Y SEXO

La tabla 49 nos transmite la clave para interpretar la representación gráfica del círculo de correlaciones de la figura 60 y muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con tipo de población que son utilizados para el análisis estadístico.

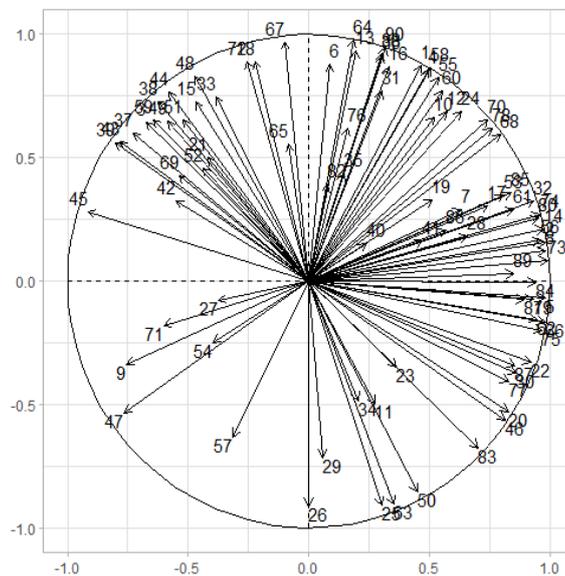


Figura 60. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave de la tabla 49.

Resultados

Tabla 49. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el tipo de población que son utilizados para el análisis estadístico e incorpora cifras de prevalencia.

Ambarroozar						Total					Mujeror						
Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
4,392593	4,489922	4,461459	4,179822	4,685406	3,950035	3,736262	4,08641	3,987742	3,380355	3,719619	3,196161	5,018262	4,891483	4,909471	4,949151	5,592694	4,638938
Ambarroozar						15-24 años					Mujeror						
Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
4,729988	5,117229	4,237719	2,956973	5,664371	5,274744	4,162501	5,203225	4,204016	3,298591	4,631805	3,091728	5,323815	5,020879	4,233813	2,577351	6,702949	7,829918
Ambarroozar						25-44 años					Mujeror						
Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
4,348045	4,320454	4,286827	4,415485	4,768317	3,88182	3,927561	4,186064	3,913493	4,087963	4,106064	3,328059	4,784133	4,458933	4,670302	4,750575	5,459145	4,476378
Ambarroozar						45-64 años					Mujeror						
Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
3,558759	3,178876	3,690959	3,578863	3,762547	3,536187	2,609066	2,754266	2,930864	1,531198	2,575961	2,805271	4,485656	3,646634	4,42598	5,558949	4,882954	4,215889
Ambarroozar						65 y más años					Mujeror						
Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor	Total	Harta 10.000 habitantor	De 10.001a 50.000 habitantor	De 50.001a 100.000 habitantor	De 100.001a 400.000 habitantor	Már do 400.000 habitantor
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
5,578479	6,124168	6,452943	5,404495	5,463302	3,991798	5,100075	5,501532	6,272934	5,02148	4,409443	3,852292	5,945602	6,649707	6,575119	5,733328	6,266254	4,111743

En el círculo de correlaciones de la figura 60 es evidente la diferencia de distribución que veíamos con respecto a la alergia. La longitud de los vectores no es similar por lo que las variables muestran diferencia entre ellas. Las variables están poco correlacionadas y se distribuyen en abanico porque existe un mayor ángulo entre ellas que el caso de la alergia. La longitud de los vectores tampoco es similar por lo que las variables muestran diferencias entre ellas.

En el cuadrante inferior izquierdo se sitúan el nicho 9 (total varones de 10.000-50.000 habitantes), el 21 (cifra total en poblaciones de 10.000-50.000 habitantes en nicho de 15-24 años), el 27 (varones de 15-24 años en poblaciones de 10.000-50.000 habitantes), el 54 (mujeres 25-44 años en poblaciones de más de 400.000 habitantes), el 57 (poblaciones de 10.000-50.000 habitantes en ambos sexos de 45-64 años) y el 71 (mujeres 45-64 años en más de 400.000 habitantes).

Todos estos nichos mantienen prevalencias alrededor del 4%, que no son de las más bajas como 2% ni de las más altas 5%-7%. En cifras totales, en poblaciones de más de 400.000 habitantes se dan las cifras más bajas en ambos sexos y también en varones. Las cifras más altas en mujeres son en poblaciones de 100.000 a 400.000 habitantes. En el caso de varones las cifras más altas corresponden a poblaciones de hasta 10.000 habitantes.

4.3.6. Análisis de componentes principales de la prevalencia de asma según nivel de estudios

- ANÁLISIS DE COMPONENTES PRINCIPALES DE LA PREVALENCIA DE ASMA POR ESTUDIOS, EDAD Y SEXO

La tabla 50 nos transmite la clave para interpretar la representación gráfica del círculo de correlaciones de la figura 61 y muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el nivel de estudios que son utilizados para el análisis estadístico.

Resultados

Tabla 50. Muestra ordenados numéricamente los diferentes nichos de edad y sexo junto con el nivel de estudios que son utilizados para el análisis estadístico e incluye la prevalencia.

Total																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
4,392,983	6,766,739	4,500,798	4,192,566	3,927,033	3,845,483	3,736,262	6,169,868	4,265,442	3,565,132	2,994,068	3,499,864	5,016,262	7,177,829	4,714,062	4,887,524	4,901,378	4,152,916
15-24 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
4,729,983	0,308,197	5,590,777	5,100,755	4,001,624	5,218,515	4,162,507	0	6,012,836	4,674,328	2,996,119	2,195,801	5,323,815	0	5,093,839	5,539,941	4,974,692	5,723,501
25-44 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
4,348,045	4,014,757	4,118,418	4,281,44	4,561,213	4,295,451	3,927,569	2,995,266	4,259,759	3,875,757	3,651,242	4,191,537	4,784,135	5,399,745	3,954,464	4,732,879	5,511,964	4,374,752
45-64 años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
3,558,792	5,358,015	4,007,282	3,512,538	3,041,782	2,997,816	2,609,656	3,138,688	3,158,898	2,583,618	2,174,097	2,697,327	4,485,658	7,061,518	4,815,779	4,469,637	3,970,192	3,309,075
65 y más años																	
Ambos sexos						Varones						Mujeres					
Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado	Total	Sin estudios	Primer grado	Segundo grado, 1º ciclo	Segundo grado, 2º ciclo	Tercer grado
5,578,479	7,717,587	4,961,615	3,911,906	3,120,736	3,182,426	5,100,749	8,224,791	4,832,982	2,854,621	2,166,125	2,483,568	5,945,601	7,454,812	5,064,132	4,839,678	4,220,465	4,295,451

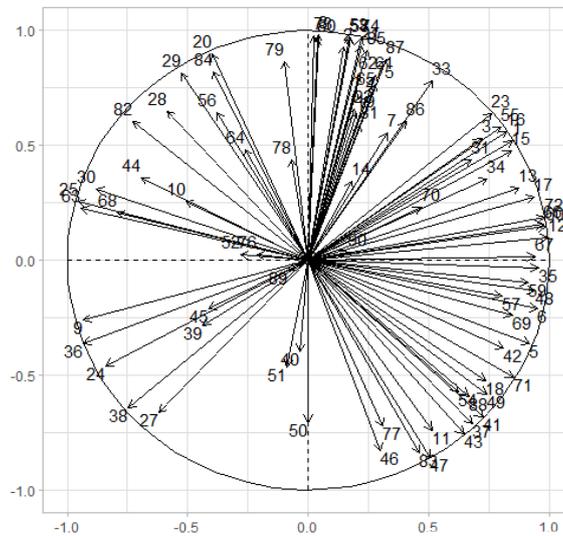


Figura 61. Muestra gráficamente mediante un biplot círculo de correlaciones las distintas variables representadas en la clave.

En figura 61 vemos el círculo de correlaciones y es evidente la diferente distribución con respecto a la alergia. Las variables están poco correlacionadas y se distribuyen en abanico porque existe un mayor ángulo entre ellas. La longitud de los vectores no es similar por lo que las variables muestran diferencia entre ellas, y es lo más marcado del gráfico, viendo que la longitud de los vectores es proporcional a la prevalencia.

Sin embargo, la distribución en los cuadrantes del círculo no nos da información relevante. En el cuadrante inferior izquierdo, el nicho 9 (total varones 1º grado 4,26%), el 24 (ambos sexos 15-24 años 3º grado 5,21%), el 27 (varones 15-24 años 1º grado 6,01%), el 36 (nicho de 15-24 años mujeres 3º grado 5,72%), y también el 38 (ambos sexos sin estudios 25-44 años 4,01%). Con vector corto 39, 45, 51 que son todos los de 1º grado de 25-44 años pero que no se asemejan en sus prevalencias y tampoco con las más bajas.

Si tenemos en cuenta el nivel de estudios clasificándolo en: sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado; vemos que en cifras totales el grupo de sin estudios tiene prevalencias más altas. En mujeres sin estudios el aumento es de un 30% más. Hay un marcado aumento en la población sin estudios, tanto de ambos sexos como por separado.

4.4. ACTIVIDAD FÍSICA EN RELACIÓN CON ALERGIA Y ASMA

4.4.1. Actividad física en relación con alergia y asma de 2011 en comunidades autónomas

Se trataba de relacionar la actividad física tanto en el trabajo como en el ocio, así como el sedentarismo con la alergia y el asma. Para ello, tomamos variables relacionadas con estos aspectos y como observaciones tomamos todas y cada una de las comunidades autónomas para las cuales existen datos cruzados de las variables reseñadas. Los datos se tomaron de dos momentos: 2011 y 2017.

Tabla 51. Clave para identificar las variables utilizadas de actividad física.

Actividad física intensa
Actividad física moderada
Actividad física ligera
Sedentarismo en ocio SI
Sedentarismo en ocio NO
Trabajo sedentario
Trabajo realizado de pie
Trabajo realizado caminando
Trabajo realizado con gran esfuerzo
Asma si
Alergia si

Tabla 52. Clave de codificación para identificar a las comunidades autónomas.

1	TOTAL
2	ANDALUCÍA
3	ARAGÓN
4	ASTURIAS, PRINCIPADO DE
5	BALEARES, ISLAS
6	CANARIAS
7	CANTABRIA
8	CASTILLA Y LEÓN
9	CASTILLA-LA MANCHA
10	CATALUÑA
11	COMUNIDAD VALENCIANA
12	EXTREMADURA
13	GALICIA
14	MADRID, COMUNIDAD DE
15	MURCIA, REGIÓN DE
16	NAVARRA, COMUNIDAD FORAL DE
17	PAÍS VASCO
18	RIOJA, LA
19	CEUTA
20	MELILLA



Figura 62. Codificación en mapa de comunidades autónomas.

4.4.1.1. Estadística descriptiva de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas

La descripción de la estadística descriptiva se desarrolla a continuación para cada una de las variables.

Tabla 53. Estadística descriptiva de 2011 de las variables en relación con todas las comunidades autónomas. Estadística descriptiva de la prevalencia de las variables con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (desviación típica, coeficiente de variación).

	<u>Media</u>	<u>Desviación estándar</u>	<u>Coeficiente variación</u>	<u>Asimetría</u>	<u>Curtosis</u>
Actividad física intensa	20,9115	5,776700	0,2762451	-0,03765876	1,63258328
Actividad física ligera	43,8795	8,749960	0,1994088	-0,33176484	-0,96634291
Actividad física moderada	20,9020	6,300768	0,3014433	0,81592649	1,28479738
Alergia	10,9595	2,739513	0,2499670	-1,64646684	4,74895229
Asma	4,0500	1,464521	0,3616101	0,53998708	0,07211411
Sedentarismo en ocio NO	58,6225	7,685305	0,1310982	-0,99012937	1,58246918
Sedentarismo en ocio SI	41,3775	7,685305	0,1857363	0,99012937	1,58246918
Trabajo caminando	12,8945	4,353407	0,3376174	0,87152930	0,89574597
Trabajo de pie	48,0645	9,700228	0,2018169	1,90555187	6,57742304
Trabajo con esfuerzo	3,3805	1,471328	0,4352396	0,09871676	-0,93773754
Trabajo sentado	35,6575	8,404735	0,2357073	-1,18813159	2,66129221

El coeficiente de variación de una distribución resulta del cociente entre la desviación típica y la media de dicha distribución. Oscila entre el valor 0 y el valor 1. Cuanto más elevado menos representativo es la media. Como vemos aquí, los valores son inferiores a 0,30, lo que demuestra un buen ajuste de la media.

Las desviaciones estándar están entre un 20-30% de la media. La asimetría y la curtosis son en su mayoría positivas lo que habla de distribuciones asimétricas a la derecha de los valores centrales.

4.4.1.2. Matriz de correlaciones de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas

La matriz de correlación contiene los coeficientes de correlación de Pearson (medida de asociación lineal entre las variables). La magnitud y el signo del coeficiente determina el grado y sentido de la relación entre las variables explicativas. Cuanto mayor sea el valor del coeficiente, mayor será la relación entre las variables. Si es positivo entonces la relación entre las variables es directa, pero si el signo es negativo la relación es inversa. Cuanto más cerca al valor uno se encuentran los coeficientes de correlación implica que la relación entre las variables es más alta, mientras que los valores próximos a cero implican ausencia de la correlación (relación).

Una vez efectuada la matriz de correlaciones da el siguiente resultado:

La alergia se relaciona con asma, actividad física moderada y trabajo sentado. Relación inversa con no sedentarismo, actividad física ligera y trabajo de pie.

El asma se relaciona con alergia, actividad física moderada y trabajo con esfuerzo. Relación inversa con no sedentarismo y actividad física ligera.

Por tanto, los factores comunes en contra de alergia y asma serían el no sedentarismo y la actividad física ligera.

Tabla 54. Resumen de la matriz de correlaciones con respecto a la alergia y el asma en 2011.

ALERGIA EN 2011		ASMA EN 2011	
Actividad física intensa	-0,05466675	Actividad física intensa	-0,17810172
Actividad física ligera	-0,33241516	Actividad física ligera	-0,34508941
Actividad física moderada	0,46093232	Actividad física moderada	0,44806741
Alergia	1,00000000	Alergia	0,66731472
Asma	0,66731472	Asma	1,00000000
NO sedentarismo en ocio	-0,29180907	NO sedentarismo en ocio	-0,38229279
SI sedentarismo en ocio	0,29180907	SI sedentarismo en ocio	0,38229279
Trabajo caminando	0,15541193	Trabajo caminando	0,19139829
Trabajo de pie	-0,42425772	Trabajo de pie	-0,09640983
Trabajo con esfuerzo	0,32509535	Trabajo con esfuerzo	0,43415639
Trabajo sentado	0,35201115	Trabajo sentado	-0,06409094

4.4.1.3. Análisis de componentes principales de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas

El análisis de componentes principales es una de las técnicas de aprendizaje no supervisado, las cuales suelen aplicarse como parte del análisis exploratorio de los datos. A diferencia de los métodos de aprendizaje supervisado, donde contamos con un grupo de variables medidas sobre un conjunto de observaciones con la intención de obtener predicciones sobre una variable respuesta, en los métodos no supervisados solo contamos con un grupo de variables de las cuales nos interesa conocer o de las que queremos extraer información, por ejemplo sobre la existencia de subgrupos entre las variables u observaciones.

Una de sus aplicaciones es la reducción de la dimensionalidad (variables), perdiendo la menor cantidad de información (varianza) posible: cuando contamos con un gran número de variables cuantitativas posiblemente correlacionadas (indicativo de información redundante), el análisis de componentes principales permite reducirlas a un número menor de variables transformadas (componentes principales que expliquen gran parte de la variabilidad de los datos). Se utiliza para sacar a relucir patrones fuertes de un conjunto de datos y para que los datos sean fáciles de visualizar y explorar.

Los autovectores son los coeficientes de la ecuación de cada componente principal. Los autovalores son los cuadrados de las desviaciones estándar. La proporción de varianza es lo que explica a cada componente principal, siendo siempre la suma igual al valor uno.

Resultados

Tabla 55. Resumen de los componentes principales ligados a la alergia y el asma en 2011. Proporción de varianza explicada por los principales componentes.

	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>	<u>Comp.5</u>
Desviación estándar	1,9474298	1,3800953	1,2608663	1,1653319	1,1224841
Proporción de varianza	0,3447712	0,1731512	0,1445258	0,1234544	0,1145428
	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>	<u>Comp.5</u>
Actividad física intensa	0,1588125				
Actividad física ligera	-0,3150599				
Actividad física moderada	0,1721964				
Alergia	0,3513997	-0,32691807	-0,07110190	-0,1352074	0,2086204
Asma	0,3123266	-0,42488845	0,22294015	-0,2132258	-0,1194218
Sedentarismo en ocio NO	-0,3935560				
Sedentarismo en ocio SI	0,3935560				
Trabajo caminando	0,1741148				
Trabajo de pie	-0,3543774				
Trabajo con gran esfuerzo	0,2966189				
Trabajo sentado	0,2667956				

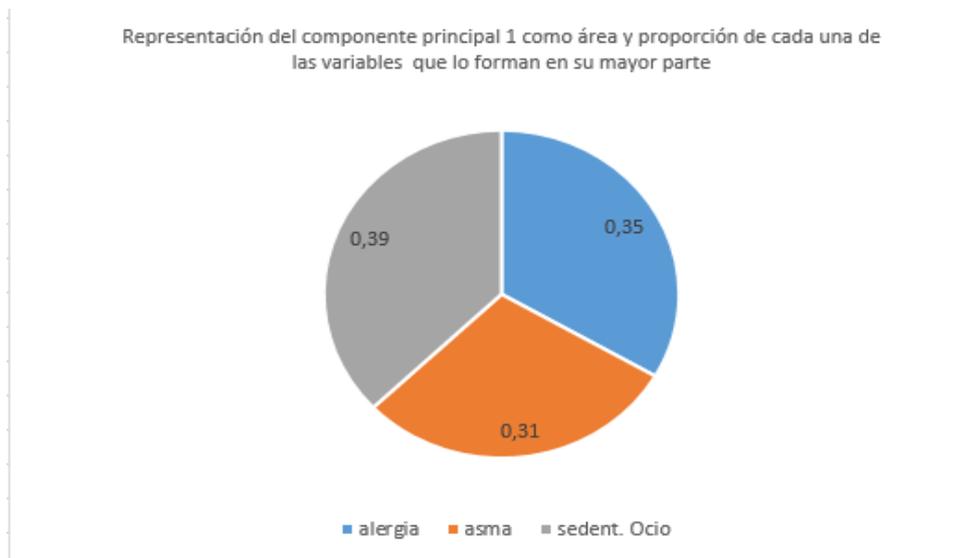


Figura 63. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.

Análisis de componentes principales de 2011 de las variables

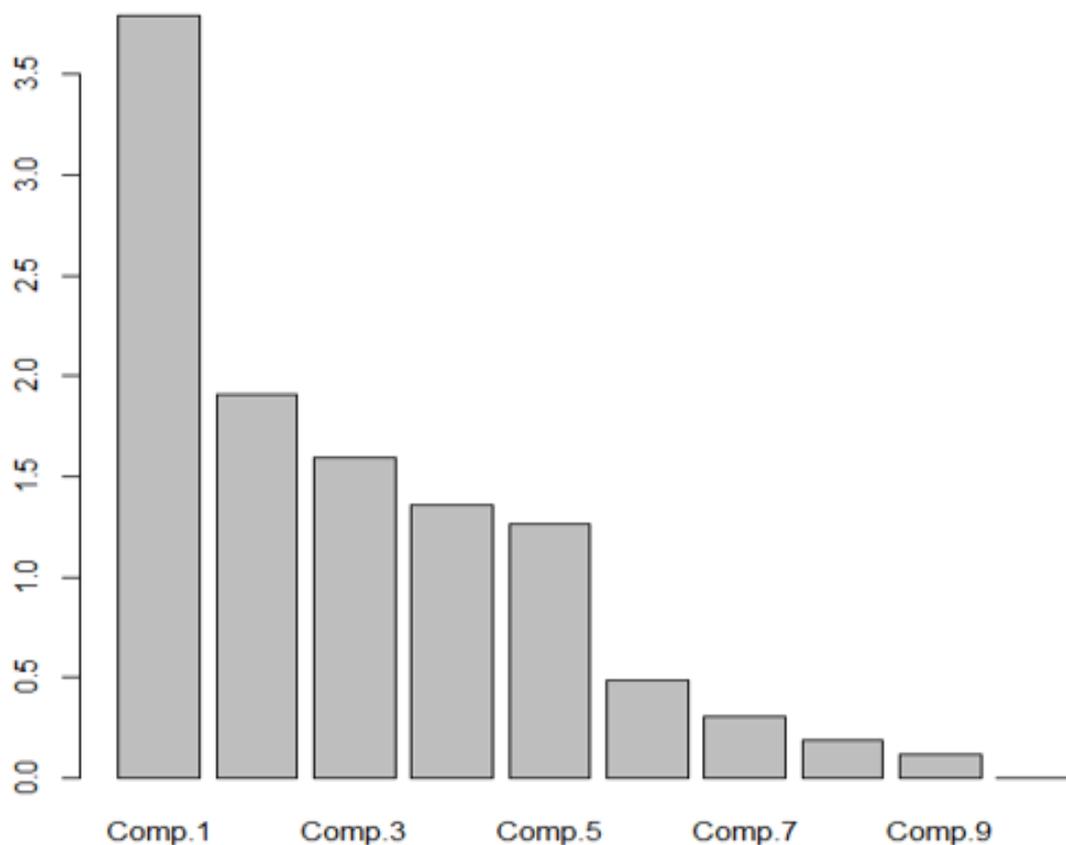


Figura 64. Gráfica de sedimentación de componentes principales de actividad física con alergia y asma. En ordenadas se expresa la varianza y en abcisas cada uno de los componentes. El asma y la alergia en 2011 se explican dentro del grupo mayoritario que corresponde al componente principal 1.

La gráfica de sedimentación muestra el número del componente principal *versus* su valor propio correspondiente y ordena los valores desde el más grande al más pequeño.

Se utiliza para seleccionar el número de componentes que se usarán con base en el tamaño de los valores propios. El patrón ideal es una curva pronunciada, seguida de una inflexión y luego de una línea recta. Hay que utilizar los componentes de la curva pronunciada antes del primer punto que indica la tendencia de línea ya que los

componentes principales restantes explican una proporción muy pequeña de la variabilidad.

El asma y la alergia en 2011 se explican dentro del grupo mayoritario que corresponde al componente principal 1. Dentro de todas las variables, el asma y la alergia estarían dentro de las que aportan más al componente 1. El sedentarismo en ocio sería el que más aportaría con 0,39 y a continuación alergia con 0,35 y asma con 0,31.

En sentido opuesto hacia el componente principal 1 estarían el no sedentarismo en ocio con -0,39, la actividad física ligera con -0,31 y el trabajo de pie con -0,35. Con una menor proporción del componente 1 estarían el trabajo con esfuerzo 0,29, el trabajo sedentario 0,26, el trabajo caminando 0,17, la actividad física intensa 0,15 y la actividad física moderada 0,17.

El componente principal 1 estaría agrupando a patologías como alergia y asma con una falta de actividad física como es el sedentarismo en ocio, al cual no se puede considerar como una patología en sí.

Es decir, el algoritmo automático del análisis de componentes principales consideraría que el sedentarismo sería o podría considerarse como una enfermedad o una situación previa a la enfermedad. En sentido opuesto hacia el componente 1 estarían el no sedentarismo en ocio y la actividad física ligera junto con el trabajo de pie.

4.4.1.4. Análisis de conglomerados de la actividad física en relación con alergia y asma de 2011 en comunidades autónomas

Mediante el algoritmo K-means del programa R obtenemos 2 *cluster* o conglomerados uno de 8 componentes (conglomerado 1) y otro de 12 (conglomerado 2).

Resultados

Tabla 56. Codificación de las comunidades autónomas con su pertenencia al conglomerado concreto.

		<u>Conglomerado</u>
1	TOTAL	2
2	ANDALUCÍA	2
3	ARAGÓN	2
4	ASTURIAS, PRINCIPADO DE	1
5	BALEARES, ISLAS	2
6	CANARIAS	1
7	CANTABRIA	2
8	CASTILLA Y LEÓN	2
9	CASTILLA-LA MANCHA	2
10	CATALUÑA	2
11	COMUNIDAD VALENCIANA	1
12	EXTREMADURA	1
13	GALICIA	2
14	MADRID, COMUNIDAD DE	2
15	MURCIA, REGIÓN DE	2
16	NAVARRA, COMUNIDAD FORAL DE	1
17	PAÍS VASCO	1
18	RIOJA, LA	1
19	CEUTA	2
20	MELILLA	1



Figura 65. Codificación en mapa de comunidades autónomas.

La Comunidad Valenciana con clave 11 se situaría en el mismo conglomerado 1 con Asturias, Navarra, La Rioja, Extremadura, Canarias y País Vasco.

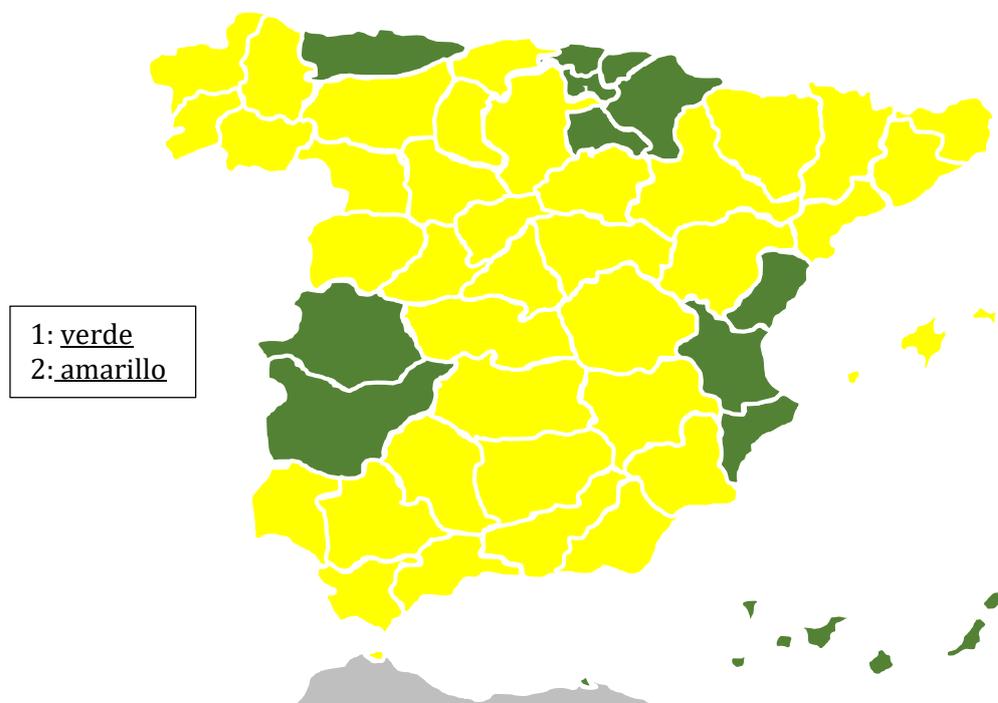


Figura 66. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color verde se muestra el conglomerado 1 y en amarillo el conglomerado 2.

Mediante el algoritmo hclust del programa R obtenemos la representación gráfica en forma de dendrograma de la distribución por conglomerados que vemos en la figura 67.

Se observa que la Comunidad Valenciana con clave 11 se situaría en el mismo conglomerado con Asturias (4), Navarra (16), La Rioja (18) y Extremadura (12) y País Vasco (17) que comparten situación en el mismo conglomerado serían una subdivisión anterior pero relacionada por la misma rama.

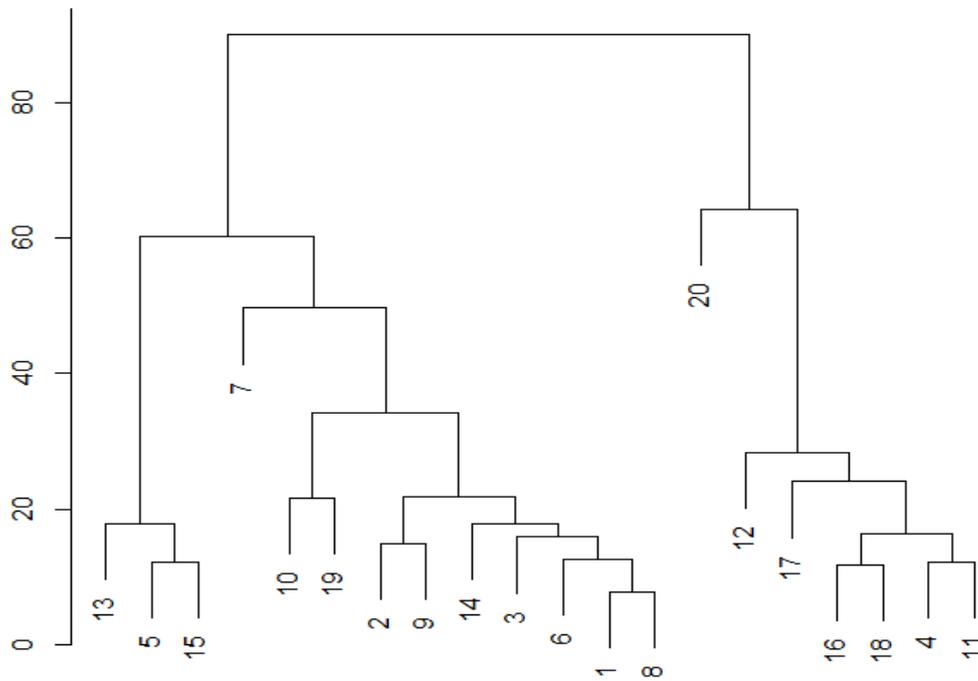


Figura 67. Dendrograma de distribución de conglomerados de comunidades autónomas en 2011 respecto a factores de salud, alergia, asma.

La gráfica de influencias dibuja los coeficientes de una variable para el primer componente *versus* los coeficientes del segundo componente. Nos permite identificar que variables tienen mayor efecto en cada componente. Las influencias pueden ir de -1 a 1. Las influencias que se aproximan a -1 y 1 indican que la variable afecta considerablemente al componente. Las influencias cercanas a 0 indican que la variable tiene poca influencia en el componente.

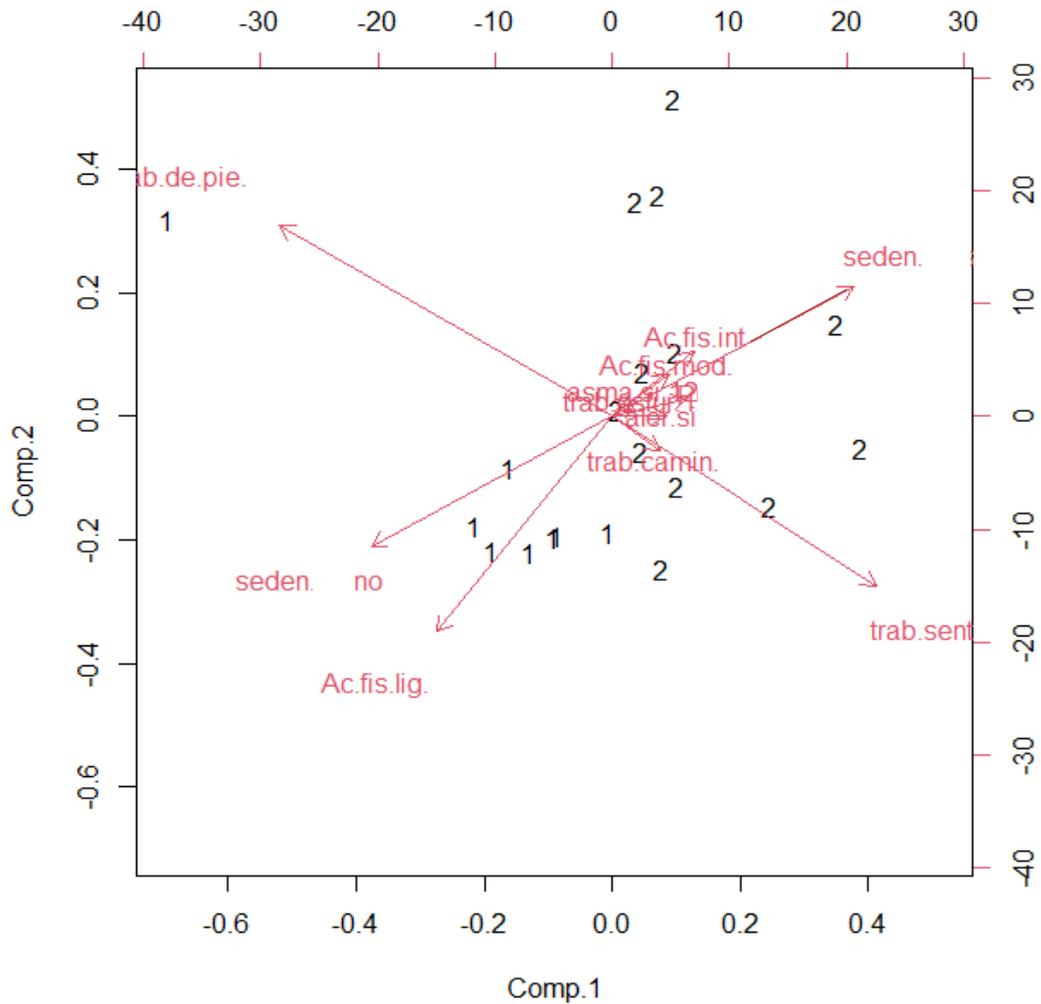


Figura 68. Gráfica de influencias biplot bidimensional de variables y *clusters* sobre componentes principales.

La abundancia de observaciones se obtiene por las proyecciones ortogonales sobre los vectores. La correlación entre variables viene dada por el ángulo existente entre vectores, siendo la mayor correlación cuando el ángulo es más pequeño y los vectores se superponen.

El sedentarismo y el trabajo sedentario tienen una correlación positiva en el eje horizontal aunque el ángulo entre ellas es grande y no existe una correlación total de la varianza. La correlación opuesta sería por parte del no sedentarismo y de la actividad física ligera.

El trabajo de pie y el trabajo sedentario se sitúan en extremos opuestos. Lo mismo con sedentarismo si y sedentarismo no.

El asma y la alergia se sitúan en la zona central alrededor del valor 0 sin relación con las variables dominantes.

Por la longitud de los vectores de las variables vemos que se confirma el sedentarismo, tanto en su versión negativa como positiva, junto con la actividad física ligera como los principales factores con influencia como factores de salud.

Las observaciones pertenecen minoritariamente al conglomerado 1 (Comunidad Valenciana, Navarra, Asturias, Navarra, La Rioja, Extremadura, Canarias y País Vasco) y mayoritariamente (resto de las comunidades autónomas) al conglomerado 2.

Por tanto, aquí se repetirían los resultados que nos dicen que el no sedentarismo y la actividad física ligera serían fundamentales en la oposición a alergia y asma, y la Comunidad Valenciana perteneciente al conglomerado 1 estaría entre las más cercanas a estos factores saludables (como se aprecia en la gráfica de influencias).

4.4.2. Actividad física en relación con alergia y asma de 2017 en comunidades autónomas

4.4.2.1. Estadística descriptiva de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas

La descripción de la estadística descriptiva se desarrolla a continuación.

Resultados

Tabla 57. Estadística descriptiva de 2017 de la prevalencia de las variables en relación con todas las comunidades autónomas. Estadística descriptiva con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (desviación típica, coeficiente de variación).

	<u>Media</u>	<u>Desviación st.</u>	<u>Coeficiente de variación</u>	<u>Asimetría</u>	<u>Curtosis</u>
Actividad física intensa	24,2810	13,039428	0,567800	1,6234516	0,9946915
Actividad física moderada	39,5220	10,928711	0,2765222	-0,66651233	0,7350694
Actividad física ligera	36,2083	16,025700	0,2105000	0,44250000	0,6749000
Alergia	14,8895	3,318193	0,2228546	-0,62858804	1,0702973
Asma	4,9085	1,348197	0,2746657	1,37847247	2,6623760
SI sedentarismo ocio	37,4155	10,626005	0,2840001	0,96645720	1,5231647
NO sedentarismo ocio	62,5845	10,626005	0,1697865	-0,96645720	1,5231647
Trabajo caminando	14,4780	5,889540	0,4067924	0,85282754	2,9913076
Trabajo de pie	43,4960	8,758862	0,2013717	-0,07431217	0,2102854
Trabajo gran esfuerzo	3,7550	2,880549	0,7671236	1,28519382	1,3334739
Trabajo sentado	36,5040	5,137331	0,1407334	0,72265168	2,0647720

El coeficiente de variación de una distribución resulta del cociente entre la desviación típica y la media de dicha distribución. Oscila entre el valor 0 y el valor 1. Cuanto más elevado menos representativa es la media. Como vemos aquí, la mayoría de valores son inferiores a 0,30, lo que demuestra un buen ajuste de la media.

Las desviaciones estándar alcanzan en algún caso valores cercanos al 50% de la media, desviaciones mayores que en el caso de los anteriores valores del 2011. La asimetría y la curtosis son en su mayoría positivas, lo que habla de distribuciones asimétricas a la derecha de los valores centrales.

Resultados

4.4.2.2. Matriz de correlaciones de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas

La matriz de correlación contiene los coeficientes de correlación de Pearson (medida de asociación lineal entre las variables). La magnitud y el signo del coeficiente determina el grado y sentido de la relación entre las variables explicativas. Cuanto mayor sea el valor del coeficiente, mayor será la relación entre las variables. Si es positivo entonces la relación entre las variables es directa, pero si el signo es negativo la relación es inversa. Cuanto más cerca al valor uno se encuentran los coeficientes de correlación implica que la relación entre las variables es más alta, mientras que los valores próximos a cero implican ausencia de la correlación (relación).

Tabla 58. Resumen de la matriz de correlaciones de factores de salud y alergia y asma en 2017.

<u>ALERGIA</u>		<u>ASMA</u>	
Actividad física ligera	0,14600825	Actividad física ligera	-0,27925165
Actividad física intensa	-0,09930766	Actividad física intensa	-0,07119500
Actividad física moderada	-0,09706243	Actividad física moderada	0,49704571
Alergia si	1,00000000	Alergia si	0,02475571
Asma si	0,02475571	Asma si	1,00000000
Sedentarismo en ocio SI	0,29292140	Sedentarismo en ocio SI	-0,08395491
Sedentarismo en ocio NO	-0,29292140	Sedentarismo en ocio NO	0,08395491
Trabajo realizado caminando	0,06711988	Trabajo realizado caminando	-0,19237382
Trabajo realizado de pie	-0,13536150	Trabajo realizado de pie	0,48498840
Trabajo realizado con gran esfuerzo	-0,29150198	Trabajo realizado con gran esfuerzo	-0,33954514
Trabajo realizado sentado	0,33067586	Trabajo realizado sentado	-0,29690737

Una vez efectuada la matriz de correlaciones da el siguiente resultado:

La alergia se relaciona con trabajo sedentario y el sedentarismo en ocio. Relación inversa con no sedentarismo en ocio y trabajo realizado con gran esfuerzo.

Resultados

El asma se relaciona con trabajo de pie y actividad física moderada. Relación inversa con trabajo realizado con gran esfuerzo y trabajo sedentario.

Por tanto, en 2017 la alergia se relaciona con trabajo sedentario y el sedentarismo en ocio. Relación inversa con no sedentarismo y trabajo realizado con gran esfuerzo. El asma con trabajo de pie y actividad física moderada. Relación inversa con trabajo realizado con gran esfuerzo y trabajo sedentario.

Y en 2011, la alergia se relaciona con actividad física moderada y trabajo sedentario. Relación inversa con no sedentarismo, actividad física ligera y trabajo de pie.

El asma se relaciona con actividad física moderada y trabajo con esfuerzo. Relación inversa con no sedentarismo en ocio y actividad física ligera.

Tabla 59. Comparativa (coeficientes de correlación de Pearson) de factores de actividad física, alergia y asma.

	Si	Si	No	No
Alergia 2017	Sedentarismo	Trabajo sedentario	No sedentarismo	Trabajo gran esfuerzo
Alergia 2011	Actividad física moderada	Trabajo sedentario	No sedentarismo	Actividad física ligera
Asma 2017	Actividad física moderada	Trabajo de pie	Trabajo gran esfuerzo	Trabajo sedentario
Asma 2011	Actividad física moderada	Trabajo gran esfuerzo	No sedentarismo	Actividad física ligera

Como vemos en la tabla comparativa, el sedentarismo y la actividad física moderada tendrían una mayor afinidad por la alergia/asma, así como el trabajo sedentario.

En contra, tendrían sobre todo el no sedentarismo en ocio acompañado de una actividad física ligera y el trabajo que requiere gran esfuerzo.

4.4.2.3. Análisis de componentes principales de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas

El análisis de componentes principales es una de las técnicas de aprendizaje no supervisado, las cuales suelen aplicarse como parte del análisis exploratorio de los datos. Una de sus aplicaciones es la reducción de la dimensionalidad (variables), perdiendo la menor cantidad de información (varianza) posible: cuando contamos con un gran número de variables cuantitativas posiblemente correlacionadas (indicativo de información redundante), el análisis de componentes principales permite reducirlas a un número menor de variables transformadas (componentes principales que expliquen gran parte de la variabilidad de los datos). Se utiliza para sacar a relucir patrones fuertes de un conjunto de datos y para que los datos sean fáciles de visualizar y explorar.

Los autovectores son los coeficientes de la ecuación de cada componente principal. Los autovalores son los cuadrados de las desviaciones estándar. La proporción de varianza es lo que explica a cada componente principal, la suma es uno.

Tabla 60. Análisis de componentes principales de 2017 de las variables en todas las comunidades autónomas.

Importancia de componentes:					
	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>	<u>Comp.5</u>
<u>Desviación estándar</u>	1,8297880	1,5741659	1,4219650	1,2469865	0,80577282
<u>Proporción de varianza</u>	0,3043749	0,2252726	0,1838168	0,1413614	0,05902453
<u>Proporción acumulativa</u>	0,3043749	0,5296475	0,7134642	0,8548257	0,91385018
	<u>Comp.6</u>	<u>Comp.7</u>	<u>Comp.8</u>	<u>Comp.9</u>	
<u>Desviación estándar</u>	0,79545673	0,44718398	0,285095532	0,18342194	
<u>Proporción de varianza</u>	0,05752286	0,01817941	0,007389042	0,00305851	
<u>Proporción acumulativa</u>	0,97137304	0,98955245	0,996941488	1,00000000	

Resultados

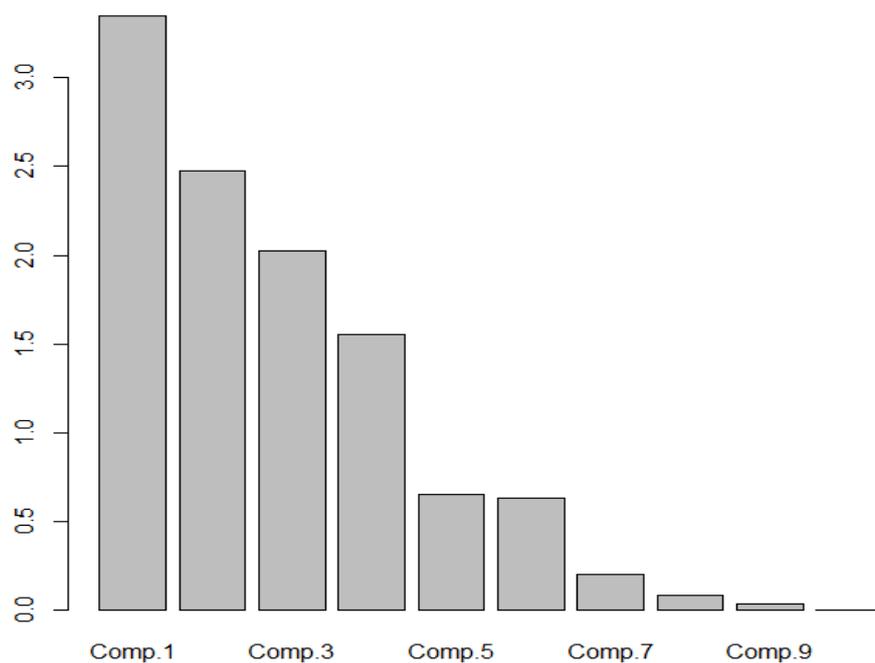


Figura 69. Gráfica de sedimentación de componentes principales de actividad física con alergia y asma en 2017.

Representación del componente principal 1 como área y proporción de cada una de las variables que lo forman en su mayor parte

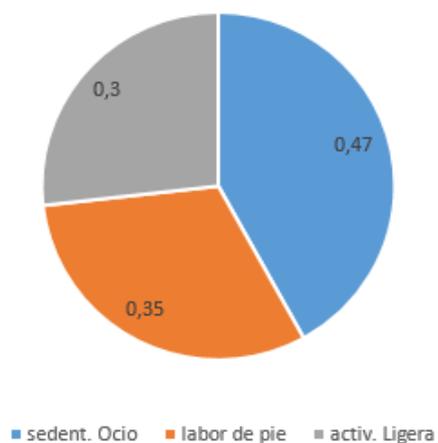


Figura 70. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.

Resultados

El asma y la alergia en 2017 no se explican o no serían lo que más influiría en el grupo mayoritario que corresponde al componente principal 1. La positividad y negatividad implica aumento proporcional o disminución proporcional con respecto al componente principal.

El componente 1 es el que comprende el mayor valor de la varianza 0,30 y afecta sobre todo a la actividad física ligera, al sedentarismo y al trabajo de pie.

La alergia se identificaría con el componente 4 y el asma con el componente 8.

Tabla 61. Resumen de los componentes principales ligados a la alergia y el asma. Proporción de varianza explicada por los principales componentes.

	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>
Actividad física ligera	<u>0,308359846</u>	0,18978475	0,52993082	0,08342002
Actividad física intensa	-0,118725113	0,12675130	-0,62074648	-0,04492930
Actividad física moderada	-0,313646595	-0,43128923	-0,04185983	-0,06950674
Alergia	-0,060595413	-0,03671139	0,13200149	0,70747999
Asma	0,001750526	-0,50308645	-0,01061896	-0,16032058
Sedentarismo en ocio	<u>0,476021926</u>	0,18849525	-0,14132981	-0,15394663
No sedentarismo en ocio	-0,476021926	-0,18849525	0,14132981	0,15394663
Trabajo caminando	-0,316426937	0,29611992	0,33088315	-0,07151731
Trabajo de pie	<u>0,357282340</u>	-0,43619534	0,01464471	-0,09817105
Trabajo de gran esfuerzo	-0,318806427	0,37670435	-0,07265484	-0,33705300
Trabajo sentado	0,078495444	0,12861298	-0,39954056	0,53310834

	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>
Alergia	-0,060595413	-0,03671139	0,13200149	<u>0,70747999</u>
Asma	0,001750526	-0,50308645	-0,01061896	-0,16032058

	<u>Comp.5</u>	<u>Comp.6</u>	<u>Comp.7</u>	<u>Comp.8</u>	<u>Comp.9</u>
Alergia	-0,44935990	0,1677418	0,20280153	-0,37696379	-0,25246056
Asma	-0,64666406	-0,2474807	0,22453151	<u>0,35030135</u>	0,26179948

El sedentarismo en ocio sería el que más aportaría con 0,47 y a continuación la actividad física ligera con 0,30 y el trabajo de pie con 0,35. En sentido negativo hacia el componente 1 estarían el NO sedentarismo en ocio con -0,47 y la actividad física moderada con -0,31. El sedentarismo en ocio sería la variable que más influiría en el conjunto de factores de salud, como también ocurría en 2011, estando este resultado en línea con la tendencia actual de considerar a esta variable como fundamental como precedente de muchas patologías. La alergia y el asma con su débil negatividad y débil positividad respectivamente serían irrelevantes como factores de salud.

4.4.2.4. Análisis de conglomerados de la actividad física en relación con alergia y asma de 2017 en comunidades autónomas

Mediante el algoritmo k-means obtenemos dos conglomerados de 18 y 2 componentes respectivamente.

Solo Murcia y Ceuta estan en el conglomerado 2, representado en la figura 71 en color azul claro.

Tabla 62. Análisis de conglomerados de 2017 de las variables en todas las comunidades autónomas.

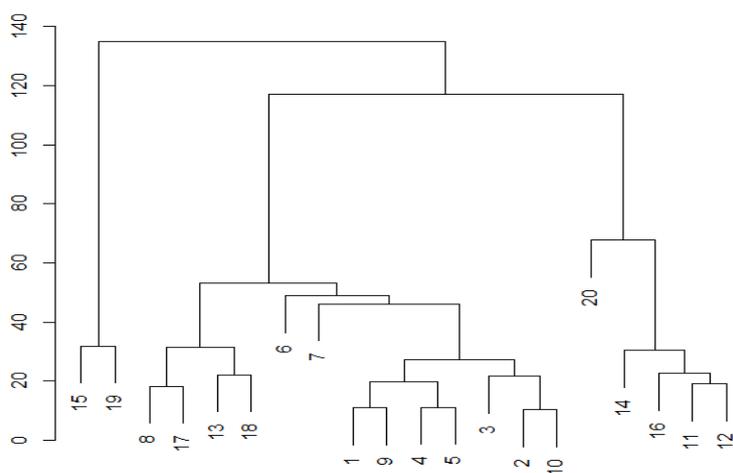
		<u>Conglomerado</u>
1	TOTAL	1
2	ANDALUCÍA	1
3	ARAGÓN	1
4	ASTURIAS, PRINCIPADO DE	1
5	BALEARES, ISLAS	1
6	CANARIAS	1
7	CANTABRIA	1
8	CASTILLA Y LEÓN	1
9	CASTILLA-LA MANCHA	1
10	CATALUÑA	1
11	COMUNIDAD VALENCIANA	1
12	EXTREMADURA	1
13	GALICIA	1
14	MADRID, COMUNIDAD DE	1
15	MURCIA, REGIÓN DE	2
16	NAVARRA, COMUNIDAD FORAL DE	1
17	PAÍS VASCO	1
18	RIOJA, LA	1
19	CEUTA	2
20	MELILLA	1



Figura 71. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color azul oscuro se muestra el conglomerado 1 y en color azul claro el conglomerado 2.

Mediante el algoritmo hclust obtenemos la representación gráfica en forma de dendrograma de la distribución por conglomerados. La Comunidad Valenciana con clave 11 se situaría con 12 Extremadura, 16 Navarra y 14 Madrid. Se volvería a repetir la asociación con Navarra y Extremadura de 2011.

Figura 72. Dendrograma de la distribución de comunidades autónomas en relación con factores de salud, alergia y asma.



La gráfica de influencias dibuja los coeficientes de una variable para el primer componente *versus* los coeficientes del segundo componente. Nos permite identificar que variables tienen mayor efecto en cada componente. Las influencias pueden ir de -1 a 1. Las influencias que se aproximan a -1 y 1 indican que la variable afecta considerablemente al componente. Las influencias cercanas a 0 indican que la variable tiene poca influencia en el componente.

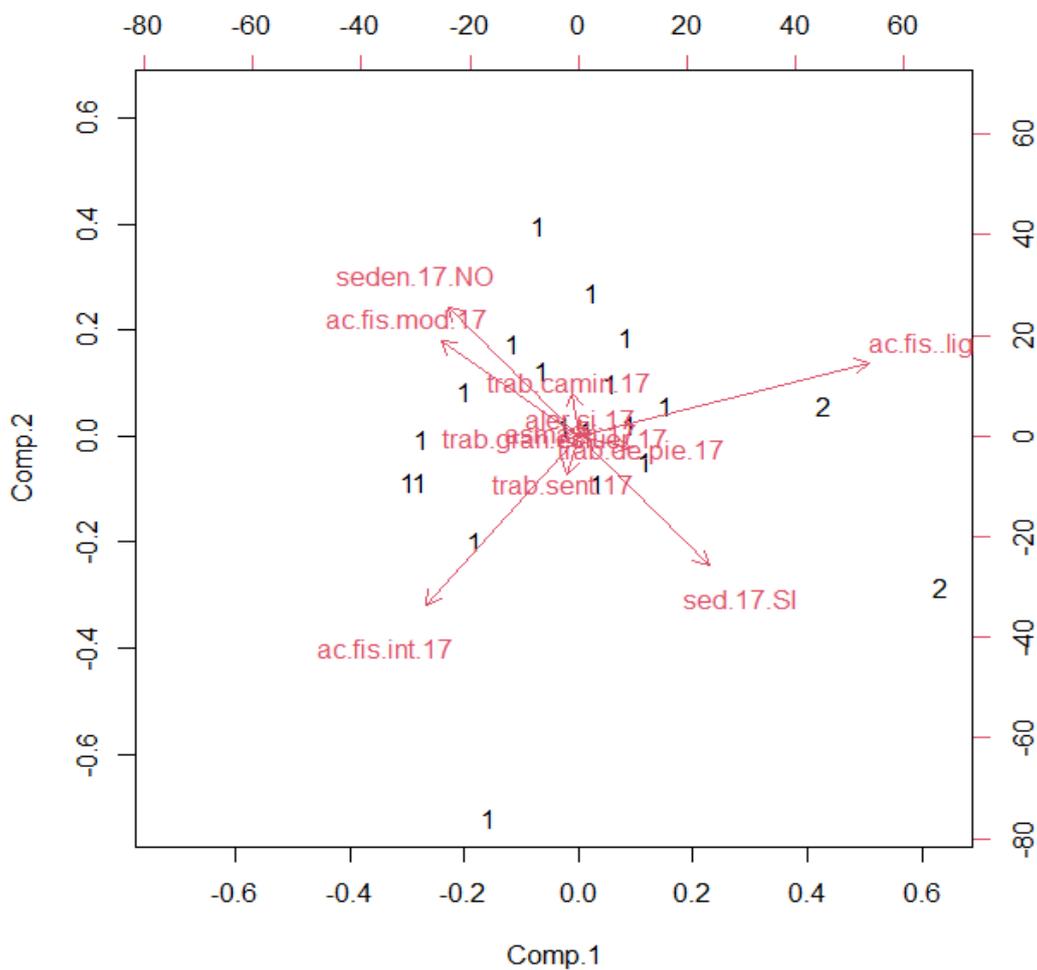


Figura 73. Gráfica de influencias biplot bidimensional de variables y conglomerados sobre los componentes principales.

La abundancia de observaciones se obtiene por las proyecciones ortogonales sobre los vectores. En el gráfico biplot no se busca la proximidad entre observaciones y variables, son las direcciones lo que importa.

La correlación entre variables viene dada por el ángulo existente entre vectores, siendo la mayor correlación cuando el ángulo es más pequeño y los vectores se superponen.

La actividad física ligera y la actividad física intensa toman extremos opuestos. Lo mismo con sedentarismo o no sedentarismo. Estas variables junto con actividad física moderada serían los vectores de mayor longitud y por tanto las más influyentes.

El asma y la alergia se sitúan en la zona central alrededor del valor 0 sin relación con las variables dominantes y con poco recorrido frente a las variables dominantes más cercanas al valor 1.

El sedentarismo y la actividad física ligera tienen una correlación positiva en el eje horizontal aunque el ángulo entre ellas es grande y no existe una correlación total de la varianza.

La correlación opuesta sería por parte del no sedentarismo en ocio y de la actividad física intensa.

Las observaciones del grupo 1 mayoritario tienen una mayor proyección ortogonal sobre las variables centrales y en el eje vertical, aproximándose más a vectores de factores saludables como es el no sedentarismo en ocio; mientras que las del grupo 2 muestran más alejamiento de las variables y más cercanía a factores no saludables como sedentarismo.

4.5. FACTORES DE SALUD Y ALERGIA-ASMA EN COMUNIDADES AUTÓNOMAS

Se trata de observar la relación entre alergia y otros factores de salud pero que no son patologías, ya que se considera que la alergia es una reacción exagerada ante agentes externos que no son nocivos para la salud pero que los pacientes sanos los identifican como tales.

Se ha escogido variables en relación con el peso corporal, hábito tabáquico, cómo se realiza el trabajo en relación con el esfuerzo físico, el sedentarismo en el ocio y el tipo de desayuno.

Se toman como observaciones cada una de las comunidades autónomas como si representaran individuos ya que existen datos agregados al respecto para todas ellas que se pueden cruzar para todas y cada una de las variables. La medición corresponde a 2017.

Tabla 63. Codificación de factores de salud con su numeración.

X1	Peso insuficiente
X2	Normopeso
X3	Sobrepeso
X4	Obesidad
X5	Fumador diario
X6	Fumador ocasional
X7	Ex-fumador
X8	Nunca ha fumado
X9	Trabajo sentado
X10	Trabajo de pie
X11	Trabajo caminando
X12	Trabajo de gran esfuerzo físico
X13	Ocio sedentario
X14	Ocio no sedentario
X15	Desayuno solo líquido
X16	Desayuno líquido y cereales/tostada
X17	Alergia
X18	Asma

Tabla 64. Referencia de las comunidades autónomas con su numeración.

1	TOTAL
2	ANDALUCÍA
3	ARAGÓN
4	ASTURIAS, PRINCIPADO DE
5	BALEARES, ISLAS
6	CANARIAS
7	CANTABRIA
8	CASTILLA Y LEÓN
9	CASTILLA-LA MANCHA
10	CATALUÑA
11	COMUNIDAD VALENCIANA
12	EXTREMADURA
13	GALICIA
14	MADRID, COMUNIDAD DE
15	MURCIA, REGIÓN DE
16	NAVARRA, COMUNIDAD FORAL DE
17	PAÍS VASCO
18	RIOJA, LA
19	CEUTA
20	MELILLA

Resultados



Figura 74. Codificación en mapa de comunidades autónomas.

4.5.1. Estadística descriptiva de factores de salud en relación con alergia y asma en comunidades autónomas

La descripción de la estadística descriptiva se desarrolla a continuación.

Tabla 65. Estadística descriptiva con valores centrales (media, mediana), de agrupación (curtosis y coeficiente de asimetría), de dispersión (coeficiente de variación, desviación típica).

	<u>Media</u>	<u>Desviación st.</u>	<u>Coefficiente variación</u>	<u>Asimetría</u>	<u>Curtosis</u>
X1 Peso insuficiente	2,1430	0,5788655	0,27011922	-0,91554916	1,40222288
X2 Normopeso	42,6780	4,2935075	0,10060236	-0,42195068	0,08403256
X3 Sobrepeso	37,5330	2,9407734	0,07835168	0,54924514	1,74465140
X4 Obesidad	17,6470	2,8631306	0,16224461	0,48457760	-0,29176647
X5 Fumador diario	22,1730	2,0326780	0,09167357	-0,71061219	-0,58728729
X6 Fumador ocasional	2,4095	0,8998273	0,37344981	0,28817241	0,82402028
X7 Ex-fumador	23,4955	5,1527514	0,21930801	-0,89380135	0,58888840
X8 Nunca ha fumado	51,9220	5,3222350	0,10250443	1,38507048	1,78188924
X9 Trabajo sentado	36,5040	5,1373313	0,14073338	0,72265168	-2,06477200
X10 Trabajo de pie	43,4960	8,7588616	0,20137166	-0,07431217	0,21028540
X11 Trabajo caminando	14,4780	5,8895399	0,40679237	0,85282754	2,99130755
X12 Trabajo de gran esfuerzo	3,7550	2,8805491	0,76712360	1,28519382	1,33347394
X13 Ocio sedentario	37,415	10,6260046	0,28400007	0,96645720	1,52316475
X14 Ocio no sedentario	62,5845	10,6260046	0,16978652	-0,96645720	1,52316475
X15 Desayuno solo líquido	9,2210	7,5986986	0,82406448	1,54084602	2,39860421
X16 Desayuno líquido/cereales	66,7615	8,8958667	0,13324846	0,07338228	0,12482418
X17 Alergia	14,8895	3,3181930	0,22285456	-0,62858804	1,07029731
X18 Asma	4,9085	1,3481967	0,666573	1,3784724	2,6623760

El coeficiente de variación de una distribución resulta del cociente entre la desviación típica y la media de dicha distribución. Oscila entre el valor 0 y el valor 1. Cuanto más elevado menos representativa es la media.

Las desviaciones estándar oscilan entre el 10-29% aunque en algún caso alcanzan valores del 50% de la media. Por los valores de asimetría y curtosis vemos que la distribución se inclina a la derecha y no es totalmente mesocúrtica.

4.5.2. Matriz de correlaciones de factores de salud en relación con alergia y asma en comunidades autónomas

La matriz de correlación contiene los coeficientes de correlación de Pearson (medida de asociación lineal entre las variables). La magnitud y el signo del coeficiente determina el grado y sentido de la relación entre las variables explicativas. Cuanto mayor sea el valor del coeficiente, mayor será la relación entre las variables. Si es positivo entonces la relación entre las variables es directa, pero si el signo es negativo la relación es inversa. Cuanto más cerca al valor uno se encuentran los coeficientes de correlación implica que la relación entre las variables es más alta, mientras que los valores próximos a cero implican ausencia de la correlación (relación).

Estableciendo la matriz de correlaciones detectamos correlación de 0,48 de x18 asma con x10 (trabajo de pie) y 0,39 con x8 (nunca ha fumado), y correlación inversa con x5 fumador diario -0,54. A su vez, x17 alergia tiene correlación de 0,33 con x9 (trabajo sedentario), de 0,38 con x7 (exfumador) y con x3 sobrepeso de 0,28; correlación inversa con x5 fumador diario -0,45.

Resultados

Tabla 66. Resumen de la matriz de correlación de 2017 de las variables en todas las comunidades autónomas.

	X1	X2	X3	X4	X5	X6
X17	-0,17101182	0,03451765	0,28656668	-0,31165564	-0,45929193	-0,09198106
X18	-0,26886916	-0,20855224	0,19342459	0,16849069	-0,54614371	-0,33978755
	X7	X8	X9	X10	X11	X12
X17	0,38937752	-0,18631061	0,33067586	-0,13536150	0,067119878	-0,29150198
X18	-0,13678253	0,39847336	-0,29690737	0,48498840	-0,192373824	-0,33954514
	X13	X14	X15	X16	X17	X18
X17	-0,29292140	0,29292140	0,02858299	-0,311657085	1,00000000	0,02475571
X18	-0,08395491	0,08395491	-0,18036138	0,122817527	0,02475571	1,00000000

4.5.3. Análisis de componentes principales de factores de salud en relación con alergia y asma en comunidades autónomas

El análisis de componentes principales es una de las técnicas de aprendizaje no supervisado. Cuando contamos con un gran número de variables cuantitativas posiblemente correlacionadas (indicativo de información redundante), el análisis de componentes principales permite reducir las a un número menor de variables transformadas (componentes principales que expliquen gran parte de la variabilidad de los datos. Se utiliza para sacar a relucir patrones fuertes de un conjunto de datos y para que los datos sean fáciles de visualizar y explorar.

Los autovectores son los coeficientes de la ecuación de cada componente principal. Los autovalores son los cuadrados de las desviaciones estándar. La proporción de la varianza es lo que explica a cada componente principal, la suma es uno.

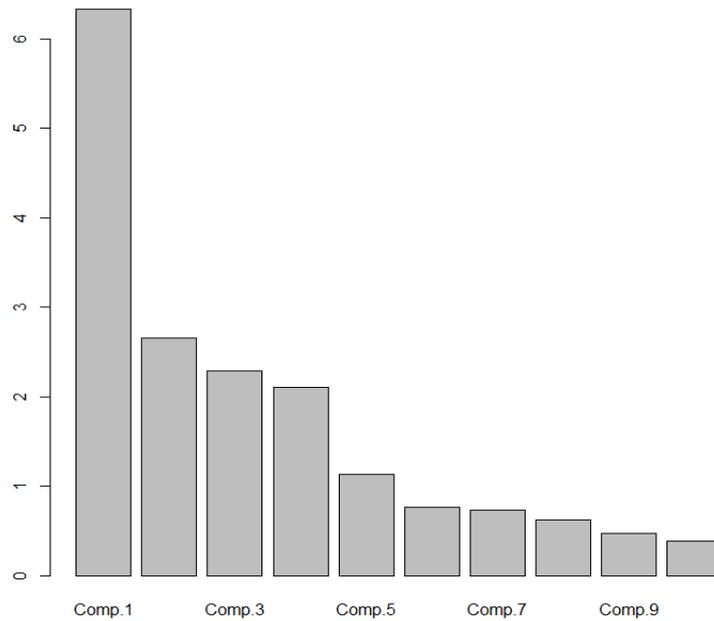
Resultados

Tabla 67. Tabla de análisis de componentes principales de factores de salud (peso, hábito tabáquico, actividad física en trabajo, sedentarismo en ocio, tipo de desayuno).

<u>ANÁLISIS DE COMPONENTES PRINCIPALES</u>						
Autovectores:						
	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>	<u>Comp.5</u>	<u>Comp.6</u>
X1	0,21940783	0,15849132	0,220584164	0,19738066	0,342260439	0,244467165
X2	0,30711798	0,03231561	0,269766234	-0,01469490	-0,289311524	0,345148444
X3	-0,21965385	-0,23125908	-0,278873442	-0,07716365	0,421123922	0,008367555
X4	-0,27902549	0,15714597	-0,163218251	0,06156361	-0,068560943	-0,576356444
X5	-0,07216981	0,49437863	-0,004177173	0,11517634	0,402570327	-0,067168412
X6	0,14256912	0,16903552	0,436163778	-0,02098025	0,004416864	-0,393464400
X7	0,35546466	-0,09537578	-0,079498932	-0,10639967	0,120079454	-0,170870496
X8	-0,34063513	-0,12500300	0,005272597	0,06278305	-0,270802288	0,257739619
X9	-0,01407062	0,07732065	0,313797589	-0,50984576	-0,250925594	-0,214189296
X10	-0,21808626	-0,26132436	0,324415794	0,26485125	0,210812261	0,095487138
X11	0,17920940	0,14468350	-0,450695770	-0,05139123	-0,102808375	0,038530456
X12	0,16385562	0,25121447	-0,396046354	0,01098208	-0,244155667	0,113108275
X13	-0,34444123	0,21867754	0,069070766	-0,05209669	-0,134871852	0,062888129
X14	0,34444123	-0,21867754	-0,069070766	0,05209669	0,134871852	-0,062888129
X15	-0,32070249	0,13023210	0,019363389	-0,26016843	0,151617272	0,240363816
X16	0,06104817	0,06839716	0,036635714	0,47178793	-0,053123451	-0,228295084
X17	0,07120058	-0,33724183	-0,027047543	-0,45936658	0,246745579	-0,116466682
X18	-0,08867686	-0,45676446	-0,043242190	0,28730554	-0,258097298	-0,182369588
	<u>Comp.7</u>	<u>Comp.8</u>	<u>Comp.9</u>	<u>Comp.10</u>	<u>Comp.11</u>	<u>Comp.12</u>
X17	-0,294885670	0,223705348	0,11419168	0,049211513	-0,060259597	0,44331905
X18	0,168048891	0,223030621	-0,11725770	0,186959270	0,172595261	-0,08467827
	<u>Comp.13</u>	<u>Comp.14</u>	<u>Comp.15</u>	<u>Comp.16</u>	<u>Comp.17</u>	
X17	0,01417371	0,31326559	0,379522974	0,0011116270	1,824402e-04	
X18	-0,63123609	0,17319731	0,084337445	0,0006745082	-2,735764e-0	
Importancia de componentes:						
	<u>Comp.1</u>	<u>Comp.2</u>	<u>Comp.3</u>	<u>Comp.4</u>	<u>Comp.5</u>	
Desviación estándar	2,5160456	1,6286426	1,5139465	1,4508557	1,06231744	
Proporción de varianza	0,3516936	0,1473598	0,1273352	0,1169435	0,06269546	
Proporción acumulativa	0,3516936	0,4990534	0,6263887	0,7433321	0,80602759	

El componente 1 que es el principal tiene poca influencia en alergia y asma. Alergia tiene más relación con el componente 12 y asma con el componente 8.

Figura 75. Gráfica de sedimentación de componentes principales.



La gráfica de sedimentación muestra que el componente 1 es el principal pero luego continúa la distribución de la curva de caída hasta el componente 5, lo que quiere decir que la importancia de la varianza de los factores de salud está repartida.

Representación del componente principal 1 como área y proporción de cada una de las variables que lo forman en su mayor parte, incluyéndose la alergia para comparar

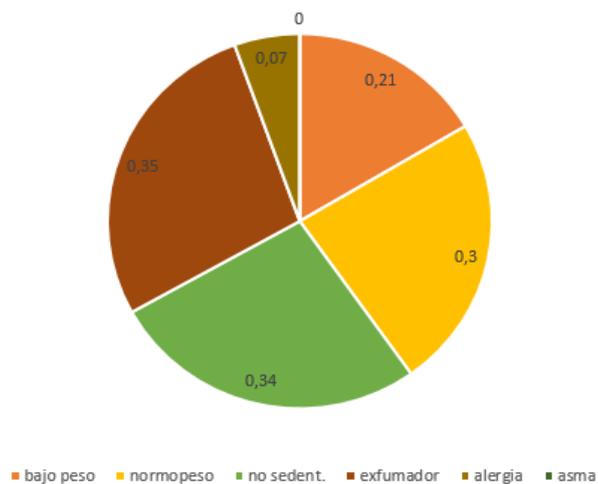


Figura 76. Representación en forma de diagrama de áreas de la proporción de las variables que forman al componente 1 principal.

Resultados

El componente 1 que es el principal y representa una proporción de la varianza de 0,35 tiene poca influencia en x17 alergia y x18 asma, y bastante en los factores de peso x1 peso insuficiente (0,21) y en x2 normopeso (0,30), también en ocio no sedentario (0,34) y ex-fumador (0,35).

Es decir, lo que formaría el estado de salud estaría muy relacionado con el normopeso o la delgadez y con el ocio no sedentario. La alergia como habitualmente se da en pacientes sanos podríamos decir que “contribuiría” o estaría representada en una mínima parte en esta teórica salud.

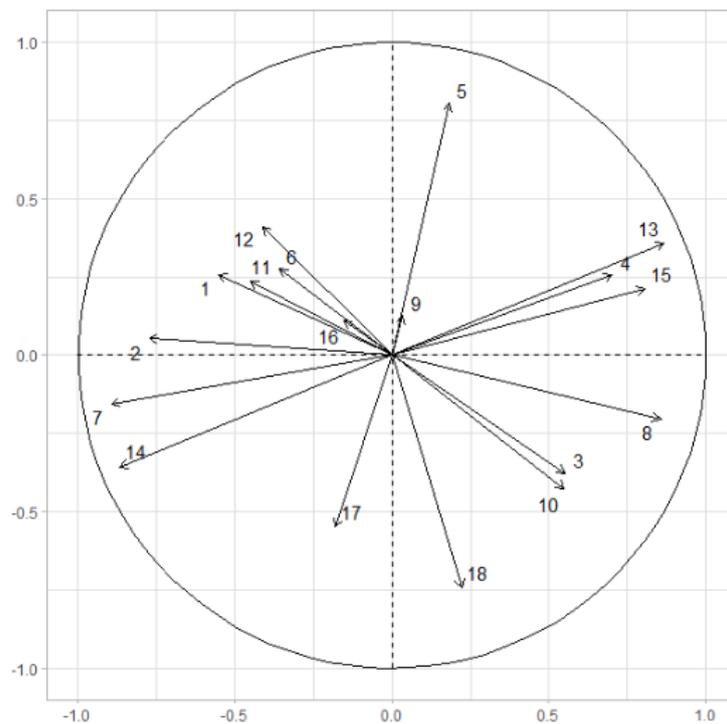


Figura 77. Círculo de correlaciones de las variables factores de salud, alergia y asma.

Los componentes principales son proyecciones que muestran donde los datos están más extendidos y presentan una mayor varianza. El círculo de correlaciones es útil para interpretar el significado de los ejes. Los vectores de cada parámetro se presentan en la gráfica frente a los dos componentes principales rotados. Cuanto más se acerca la flecha de la variable al perímetro del círculo, mayor es la significatividad de la variable.

Los que más se aproximan a -1 son x14 (no sedentarismo), x7 (ex-fumador) y x2 (normopeso). Por oposición a los anteriores factores saludables estarían los no saludables como x13 (sedentarismo en ocio), x15 (desayuno solo líquido) y x4 (obesidad) que se sitúan en el extremo contrario. Fumador x5 está en extremo opuesto a x17 alergia y x18 asma.

Los factores de peso se distribuyen en abánico a lo largo de todo el círculo de correlaciones puesto que expresan conceptos opuestos, así x1 peso insuficiente y x2 normopeso está en extremo opuesto a x3 sobrepeso y x4 obesidad. Esta última está en el extremo opuesto a x14 no sedentarismo y con un valor de longitud ligeramente inferior a x13 sedentarismo.

4.5.4. Análisis de conglomerados de factores de salud en relación con alergia y asma en comunidades autónomas

Efectuando el algoritmo K-means en el programa R obtenemos:

La columna de la derecha refleja la distribución de cada comunidad autónoma en conglomerado 1 ó en 2. La mayoría de comunidades están en el conglomerado 1 (expresado en color ocre en la gráfica). La Comunidad Valenciana está en el conglomerado 1. Murcia, Cantabria, Ceuta y Melilla forman la agrupación 2 (expresado en color naranja en la gráfica).

Resultados

Tabla 68. Tabla de codificación de comunidades autónomas con sus conglomerados correspondientes.

		<u>Conglomerado</u>
1	TOTAL	1
2	ANDALUCÍA	1
3	ARAGÓN	1
4	ASTURIAS,PRINCIPADO DE	1
5	BALEARES, ISLAS	1
6	CANARIAS	1
7	CANTABRIA	2
8	CASTILLA Y LEÓN	1
9	CASTILLA-LA MANCHA	1
10	CATALUÑA	1
11	COMUNIDAD VALENCIANA	1
12	EXTREMADURA	1
13	GALICIA	1
14	MADRID, COMUNIDAD DE	1
15	MURCIA, REGIÓN DE	2
16	NAVARRA,COMUNIDAD FORAL	1
17	PAÍS VASCO	1
18	RIOJA, LA	1
19	CEUTA	2
20	MELILLA	2



Figura 78. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados. En color ocre se muestra el conglomerado 1 y en color naranja el conglomerado 2.

La gráfica de influencias dibuja los coeficientes de una variable para el primer componente *versus* los coeficientes del segundo componente. Nos permite identificar que variables tienen mayor efecto en cada componente. Las influencias pueden ir de -1 a 1. Las influencias que se aproximan a -1 y 1 indican que la variable afecta considerablemente al componente. Las influencias cercanas a 0 indican que la variable tiene poca influencia en el componente. La correlación entre variables viene dada por el ángulo existente entre vectores, siendo la mayor correlación cuando el ángulo es más pequeño y los vectores se superponen.

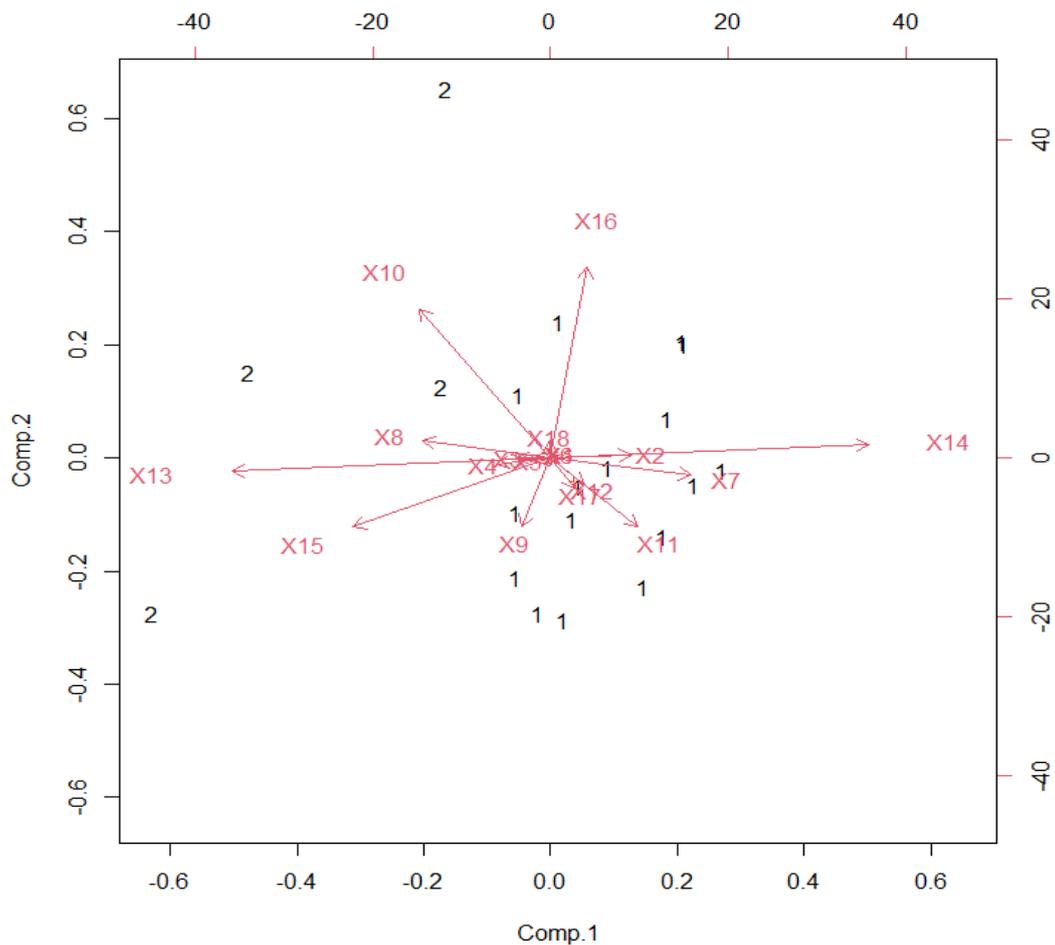


Figura 79. Gráfica de influencias biplot bidimensional de variables (factores de salud, nominados con X y numerados) y observaciones (conglomerados numerados con el 1 ó el 2) sobre componentes principales.

La gráfica muestra a las variables numeradas y su relación con los dos conglomerados. Mayoritariamente pertenecen al conglomerado 1. Las muestras del grupo 1 mayoritario tienen una mayor proyección ortogonal sobre las variables centrales y en el eje vertical, mientras que las del conglomerado 2 muestran más alejamiento de las variables. La abundancia de observaciones se obtiene por las proyecciones ortogonales sobre los vectores.

El trabajo de pie x10 y el trabajo caminando x11 se sitúan en extremos opuestos. Lo mismo ocurre con x13 sedentarismo en ocio y x14 no sedentarismo en ocio.

El no sedentarismo x14 tiene una clara oposición con el sedentarismo x13, luego el eje horizontal nos mostraría sobre todo este tipo de variables, x14 (no sedentarismo en ocio) que es el que más se aproxima a 1 y tiene influencia grande en el componente 1 y x13 (sedentarismo en ocio) tiene influencias negativas en el componente 1 con aproximación a -1.

También el x7 ex-fumador se situaría en la línea del x14 no sedentarismo aunque con una menor longitud del vector.

Tendrían poca influencia el x10 trabajo de pie y el x16 desayuno con cereales/tostadas/bollería que están en ángulo recto con el eje horizontal que integra a x17 y x18.

El asma y la alergia se sitúan en la zona central alrededor del valor 0 sin relación con las variables dominantes y con poco recorrido frente a las variables dominantes más cercanas al valor 1.

Las comunidades del conglomerado 2 son las más alejadas de los vectores con más longitud, es decir los más influyentes. Además, está sobre todo alejadas de los vectores que representan lo más saludable como x14 no sedentarismo.

Mediante el algoritmo hclust obtenemos la representación gráfica en forma de dendrograma de la distribución por conglomerados

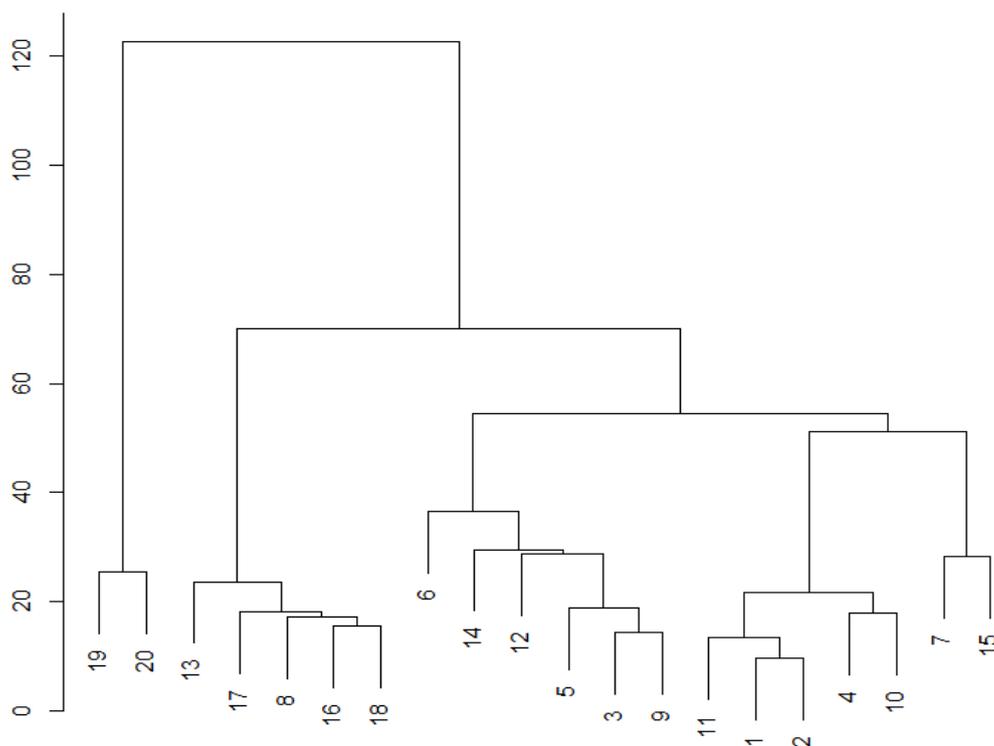


Figura 80. Dendrograma de jerarquización de comunidades autónomas en relación con factores de salud.

La Comunidad Valenciana 11 se agrupa con 1 (media total) y con 2 Andalucía, con 4 Asturias y con 10 Cataluña, y en menor medida con 7 Cantabria y con 15 Murcia. Ocupan uno de los extremos 7 Cantabria y 15 Murcia, y en el otro 19 Ceuta y 20 Melilla que formaban el conglomerado 2.

Se estaría reproduciendo el resultado de reparto de conglomerados estando en los puntos más alejados las comunidades del conglomerado 2.

La gráfica de puntuaciones (algoritmo MASS de R) representa visualmente las puntuaciones del segundo componente principal *versus* las puntuaciones del primer componente principal. Si los dos primeros componentes explican la mayor parte de

la varianza de los datos se puede usar esta gráfica para ver agrupaciones y tendencias.

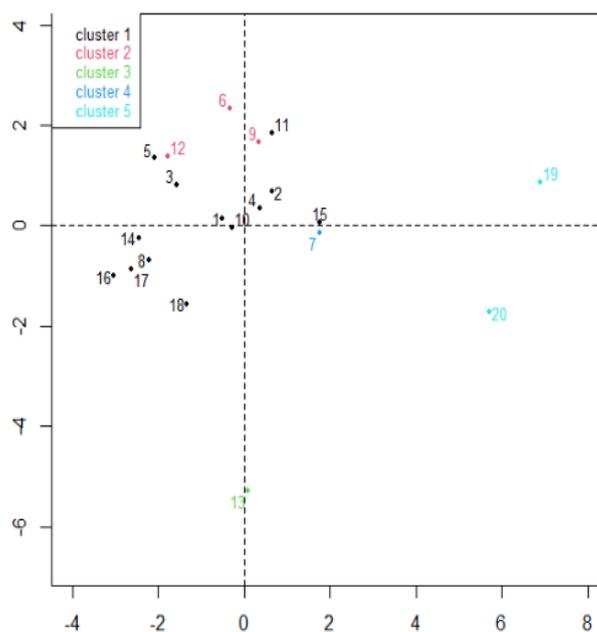


Figura 81. Gráfica de puntuaciones con algoritmo MASS. Clasificación por conglomerados de comunidades autónomas en relación con factores de salud.

Visualizamos de forma clara que tanto Ceuta y Melilla que formaban el conglomerado 2 serían los más alejados del resto de puntos y como hemos visto antes también los más alejados de factores saludables.

Resultados

Tabla 69. Tabla de codificación de comunidades autónomas con sus conglomerados correspondientes.

		<u>Conglomerado</u>
1	TOTAL	1
2	ANDALUCÍA	1
3	ARAGÓN	1
4	ASTURIAS,PRINCIPADO DE	1
5	BALEARES, ISLAS	1
6	CANARIAS	2
7	CANTABRIA	4
8	CASTILLA Y LEÓN	1
9	CASTILLA-LA MANCHA	2
10	CATALUÑA	1
11	COMUNIDAD VALENCIANA	1
12	EXTREMADURA	2
13	GALICIA	3
14	MADRID, COMUNIDAD DE	1
15	MURCIA, REGIÓN DE	2
16	NAVARRA,COMUNIDAD FORAL	1
17	PAÍS VASCO	1
18	RIOJA, LA	1
19	CEUTA	5
20	MELILLA	5

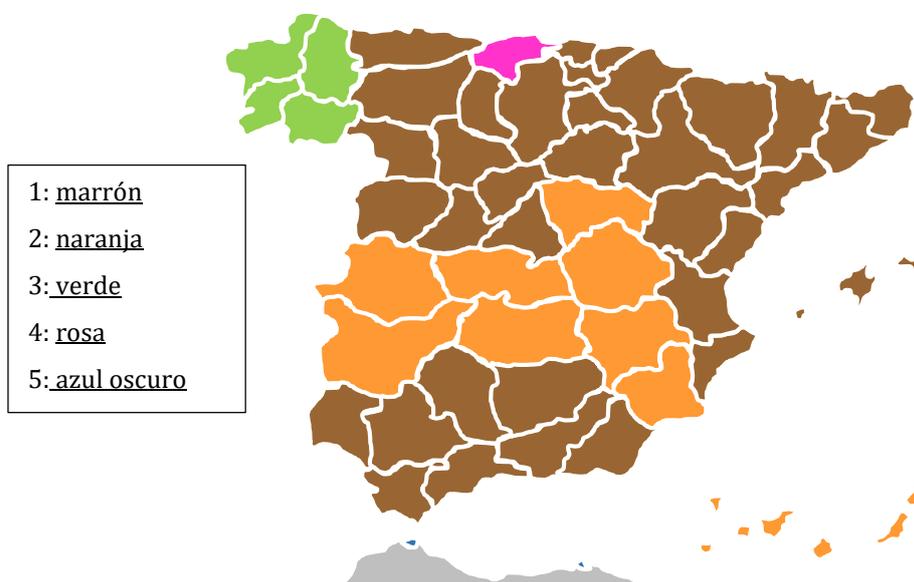


Figura 82. Muestra el mapa con las comunidades autónomas y su pertenencia a los conglomerados según los colores de la leyenda.

Podemos explicar la pertenencia a cada uno de los grupos realizando un análisis de la varianza.

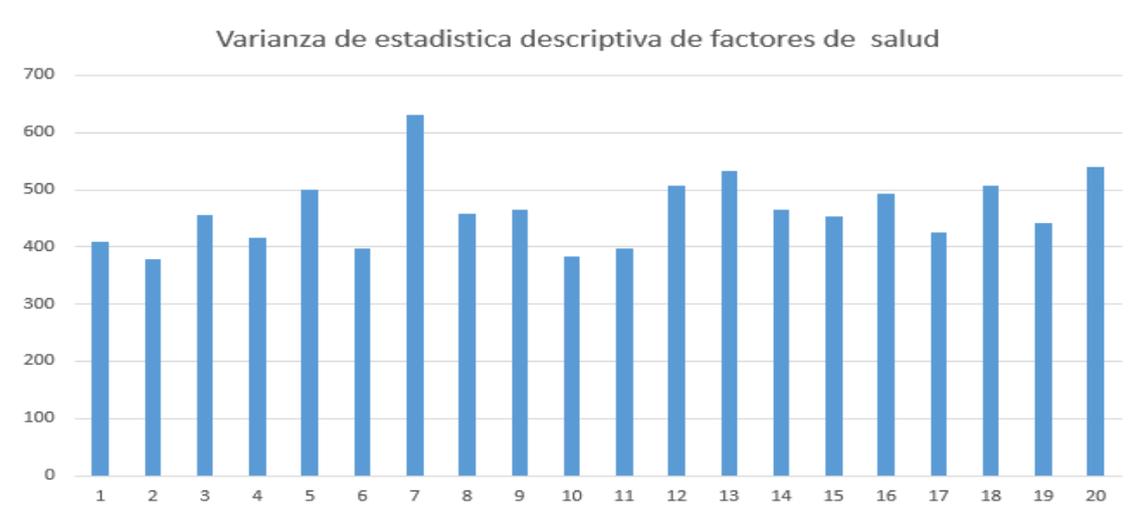


Figura 83. Varianza de estadística descriptiva de factores de salud numerada según clave de comunidades autónomas.

Asociando las figuras 81 y 83, obtendríamos que las que superarían el valor de 500 serían, 13 Galicia que es el componente más separado (y es el único del conglomerado 3), más el 7 Cantabria (también único componente del conglomerado 4) muy unido al anterior. Los miembros del conglomerado 5 (19 Ceuta, 20 Melilla) también ampliamente separados. En el caso de 19 no tendría un valor tan alto.

La varianza influiría en la distribución de conglomerados puesto que las de menor valor (no superando 400) 2, 4, 10 y 11 estarían en el conglomerado 1 y en el cuadrante superior derecho. Es decir, la Comunidad Valenciana 11 se alinearía con las de menor varianza: 2 Andalucía, 4 Asturias y 10 Cataluña.

En el cuadrante superior izquierdo, con valor entre 400 y 500 de varianza, se situarían 3 Aragón, 5 Baleares, 6 Canarias y 12 Extremadura.

En cuadrante inferior izquierdo: 8 Castilla y León, 14 Madrid, 16 Navarra, 17 País Vasco, 18 La Rioja.

El componente principal que forma a la varianza en un 35% representa sobre todo los niveles de peso corporal como factor de salud, y ahí estarían las comunidades 11, 2, 4, 10 del conglomerado 1 con niveles bajos en la gráfica 83.

La distribución mantiene semejanzas parciales con el dendrograma pero en el gráfico de puntuaciones se representan las observaciones o marcas en los ejes formados por los dos primeros componentes principales.

4.6. COMPARATIVA DE LA PREVALENCIA DE ALERGIA Y ASMA POR COMUNIDADES AUTÓNOMAS

Uno de nuestros objetivos era determinar la prevalencia de alergia y asma pero diseccionando también datos por comunidades autónomas, y en especial los referidos a la Comunidad Valenciana.

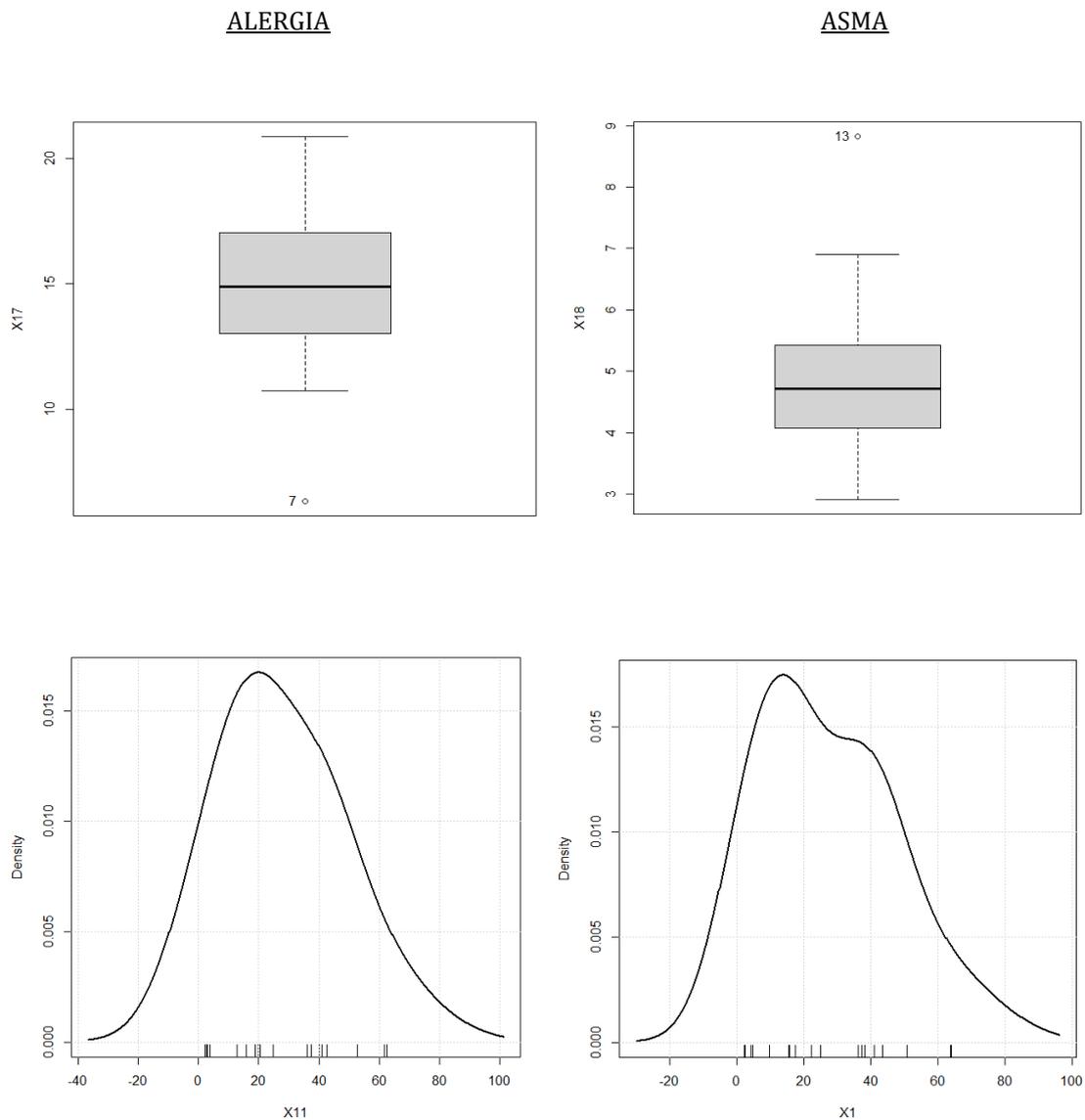
Hemos efectuado un análisis de la prevalencia de alergia y asma en las comunidades autónomas.

En primer lugar mostramos mediante diagramas de caja la distribución estadística de la alergia y el asma a nivel nacional. La figura 84 nos visualiza una distribución muy regular con una mediana muy centrada en ambos casos y con escasos valores atípicos, aunque sí con valor máximo y mínimo separados a distancia del resto.

La comparación de los datos, mediante el diagrama de densidad de la figura 84, entre la comunidad Valenciana (x11) y el total nacional (x1) nos revela una distribución muy normalizada en la Comunidad Valenciana acercándose en gran medida a los datos nacionales.

Resultados

Figura 84. La figura muestra la distribución de la prevalencia de alergia y asma a nivel nacional mediante diagrama de caja y diagrama de densidad. En el caso de alergia el diagrama de densidad muestra claramente la normalidad de la distribución y el diagrama de cajas corrobora esta distribución normal con un único valor atípico. En el caso del asma la distribución normal también se confirma con una ligera irregularidad.



El diagrama de líneas de la figura 85 nos revela una proporción de alergia y asma en la Comunidad Valenciana muy en términos medios no destacando ni por su elevada prevalencia ni por su escasa prevalencia.

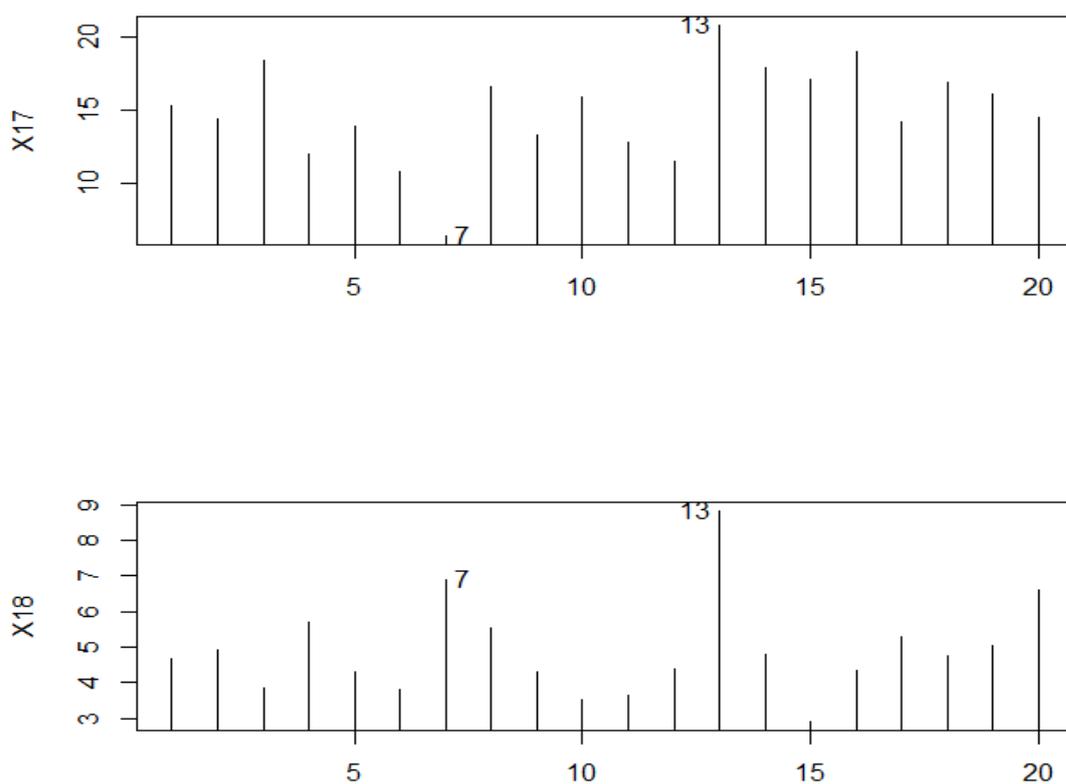


Figura 85. Diagrama de líneas mostrando en abcisas las comunidades autónomas numeradas de 1 a 20 y en ordenadas la prevalencia de alergia (X17) y de asma (X18).

Las figuras 87 y 88 nos dibujan en forma de diagrama de barras tanto la prevalencia de alergia y asma en 2011 como en 2017 en todas y cada una de las comunidades autónomas. También se expresa el valor numérico en ellas para poder hacer una comparativa mejor.

Resultados



Figura 86. Codificación en mapa de comunidades autónomas.

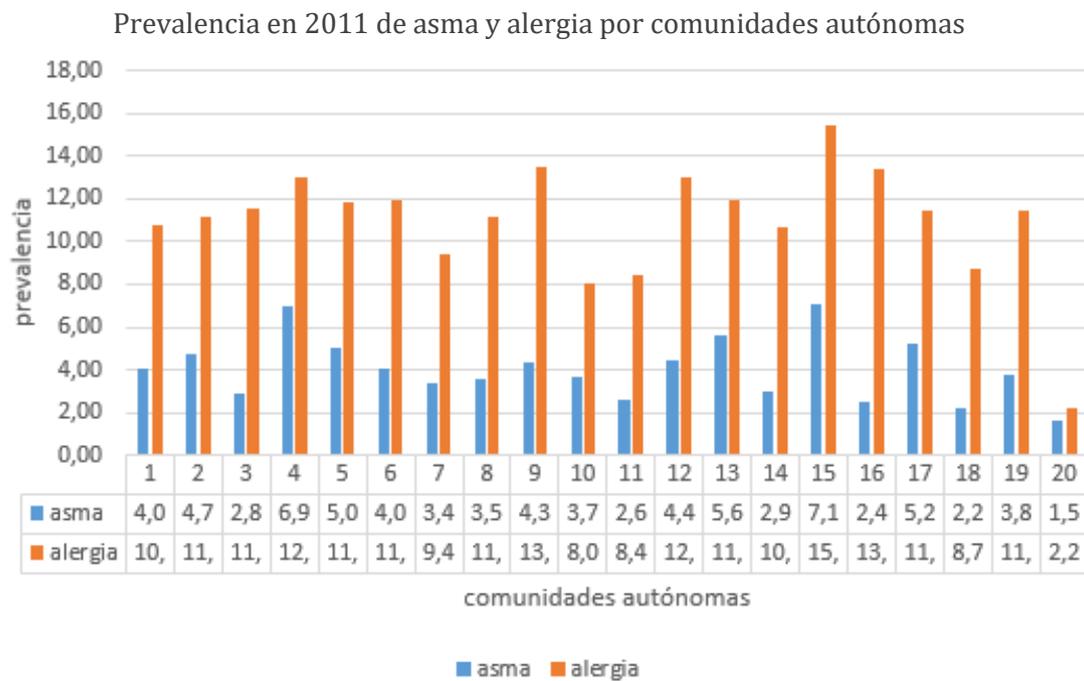


Figura 87. Diagrama de barras mostrando en abcisas las comunidades autónomas numeradas de 1 a 20, junto con la prevalencia de alergia y asma de cada una de ellas en 2011.

Resultados



Figura 88. Diagrama de barras mostrando en abcisas las comunidades autónomas numeradas de 1 a 20, junto con la prevalencia de alergia y asma de cada una de ellas en 2017.

En general, la mayoría de comunidades autónomas muestra un crecimiento de alergia:

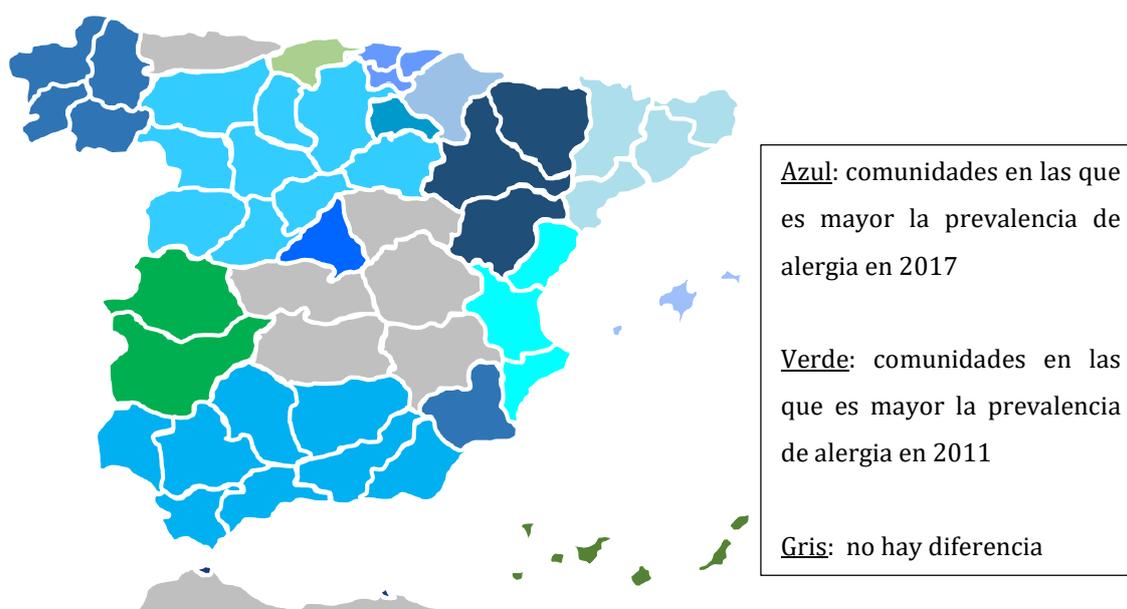


Figura 89. Mapa de prevalencia de alergia por comunidades autónomas según color de la leyenda.

La figura 89 nos muestra la situación siguiente:

Las comunidades autónomas como Galicia, País Vasco, La Rioja, Navarra, Madrid, Castilla-León, Ceuta y Melilla, Aragón, Cataluña, Comunidad Valenciana, Baleares, Murcia y Andalucía, tienen más alergia en 2017 que en 2011. Extremadura, Cantabria y Canarias tienen más alergia en 2011. En cambio, Asturias y Castilla-La Mancha se mantiene en cifras similares.

En 2011, era Murcia con un 15% quien lideraba la prevalencia de alergia. Entre el 10-15% estaban la mayoría de comunidades: Andalucía, Aragón, Cantabria, Baleares, Canarias, Castilla-León, Castilla-La Mancha, Extremadura, Galicia, Madrid, Navarra, País Vasco y Ceuta. Y con menos del 10% estaban Cantabria, La Rioja, Cataluña y la Comunidad Valenciana. A destacar la baja cifra de Melilla con un 2,2%.

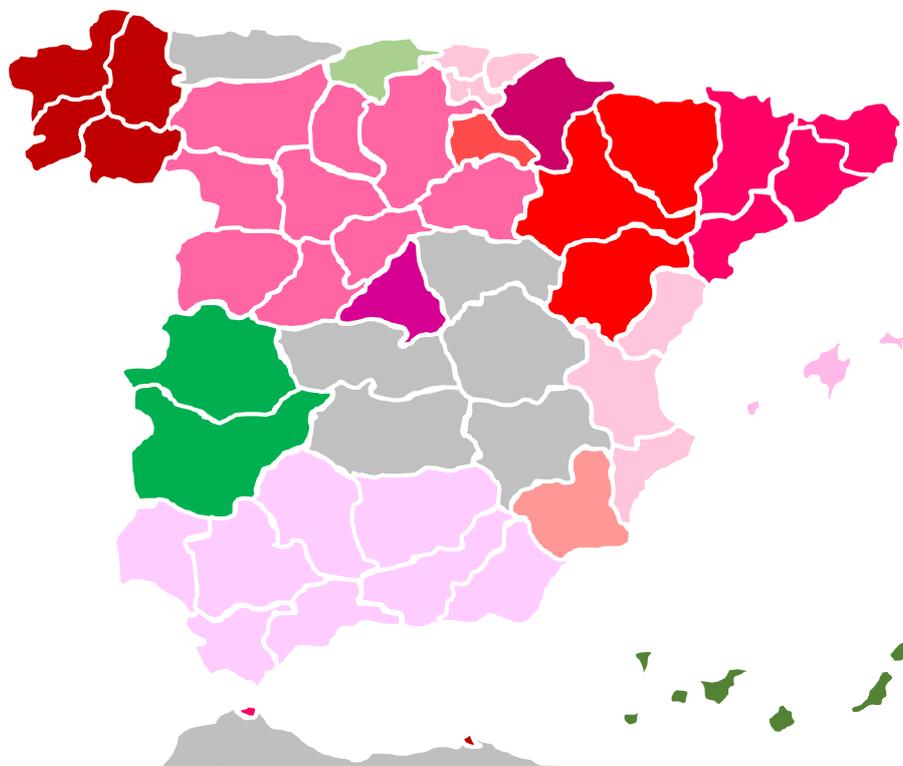
En 2017, Galicia era la comunidad con la prevalencia de alergia más importante (20%), Navarra con un 19% y Aragón con un 18% le seguían. Después aparecían una serie de comunidades con cifras superiores al 15% como Cataluña, Madrid, Castilla-León, Ceuta y Murcia. Con cifras inferiores al 15% una buena parte de las comunidades: País Vasco, Castilla-La Mancha, Extremadura, Canarias, Baleares, Asturias, Andalucía y Melilla. También la Comunidad Valenciana estaría en este grupo con un 11%. A destacar la baja prevalencia en Cantabria con un 6,3%.

Pero el aumento no es el mismo en todos los casos. Es muy llamativo el crecimiento en Melilla, del 2,2% al 14% y de Galicia de 11% al 20%.

De la misma forma con una elevación importante: Aragón que ha pasado de un 11% a un 18% y Navarra de un 13% a un 19%, Cataluña de un 8% a un 15%, Madrid de un 10% a un 17%, La Rioja de un 8,7% a un 16%, Ceuta de un 11% a un 16% y Castilla-León de un 11% a un 16%. Por otro lado, con un crecimiento más moderado: Andalucía del 11% al 14%, País Vasco del 11% al 14%, Baleares de 11% a 13% y Murcia de 15% a 17%. La Comunidad Valenciana ha pasado de un 8,4% a un 12%.

Por tanto, como vemos en la figura 90, la mayoría de comunidades autónomas han aumentado en prevalencia de alergia: Galicia, País Vasco, La Rioja, Navarra, Madrid,

Castilla-León, Ceuta y Melilla, Aragón, Baleares, Cataluña, Comunidad Valenciana, Murcia y Andalucía ya que tienen más alergia en 2017 que en 2011.



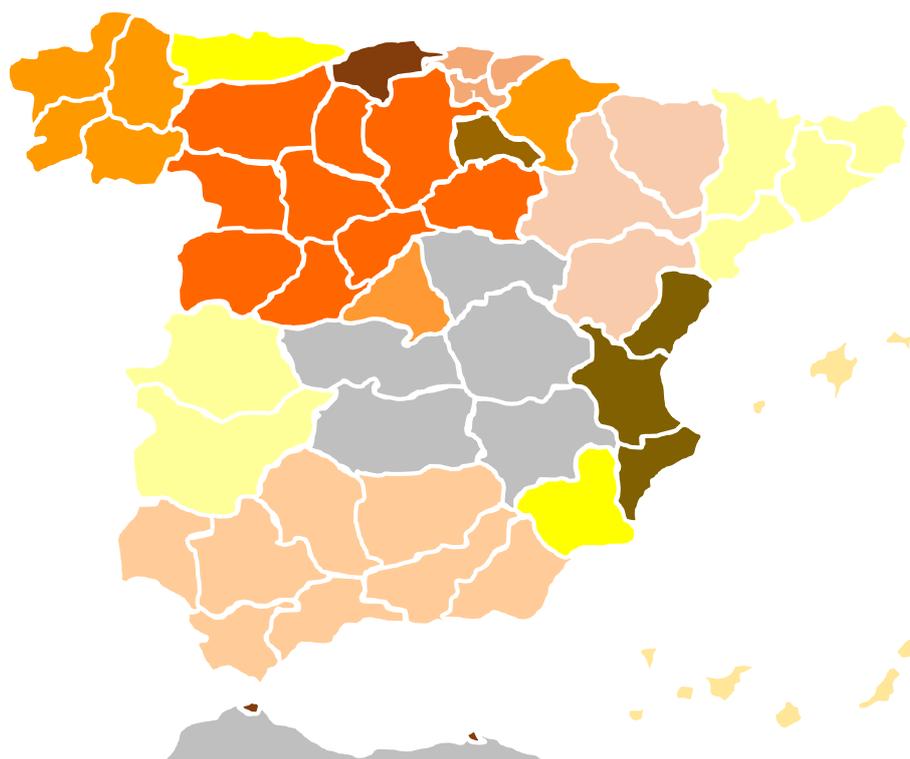
Colores con tonalidades rosas-rojo con diferentes intensidades mostrando las comunidades autónomas donde ha aumentado más y en mayor cantidad la alergia.
Colores verdes representando las comunidades autónomas en las que ha habido menos alergia.
Color gris muestra las comunidades autónomas donde no ha habido cambios en la evolución de la alergia.

Figura 90. Mapa de evolución de prevalencia de alergia en España según los colores de la leyenda.

Las comunidades autónomas donde se observa menos alergia en 2017 que en 2011 serían Extremadura (del 12% al 11%), Cantabria (del 9,4% al 6,3%) y Canarias (del 11% al 10%). Asturias se mantiene en un 12% y Castilla-La Mancha se ha mantenido en un 13%.

La figura 91 nos visualiza la siguiente situación:

Las comunidades autónomas como Galicia, Cantabria, La Rioja, Navarra, Madrid, Castilla-León, Andalucía, Ceuta y Melilla tienen más asma en 2017 que en 2011. Extremadura, Asturias, Cataluña, Murcia, Canarias y Baleares tienen más asma en 2011. En cambio, Castilla-La Mancha muestra cifras similares.



Naranja-marrón: comunidades en las que es mayor la prevalencia de asma en 2017

Amarillo: comunidades en las que es mayor la prevalencia de asma en 2011

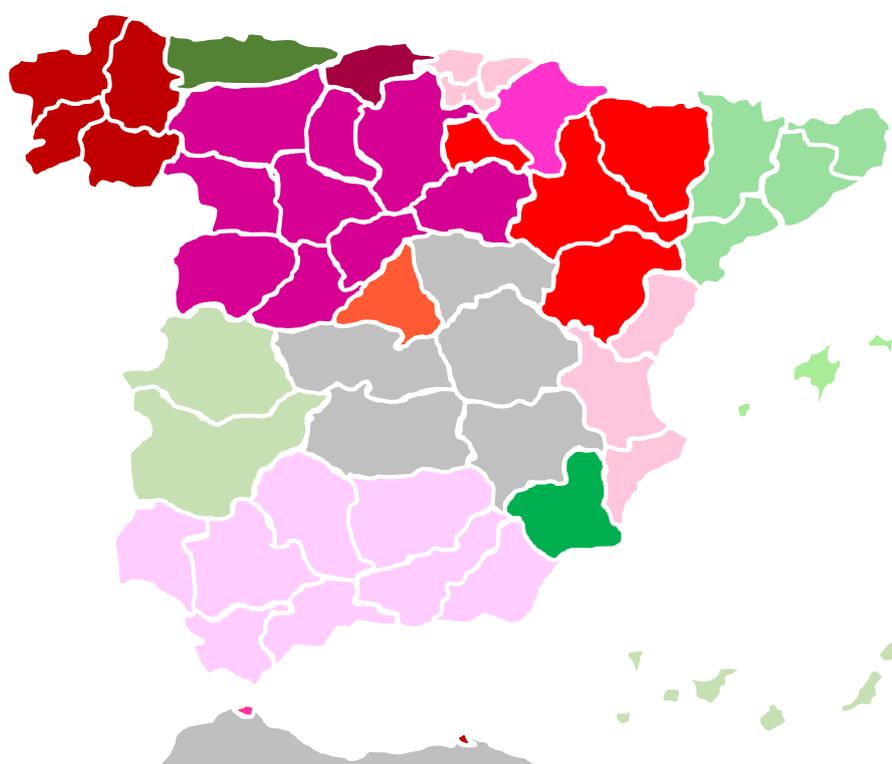
Gris: no hay diferencia

Figura 91. Mapa de prevalencia de asma en comunidades autónomas según color de la leyenda.

En lo referente a la prevalencia del asma, tienen más asma en 2017 que en 2011: Melilla (del 1,5 al 6,6%), Cantabria (del 3,4% al 6,9%), Galicia (del 5,6% al 8,8%), La Rioja (del 2,2% al 4,7%), Castilla-León (del 3,5% al 5,5%), Navarra (del 2,4 al

4,3%), Madrid (del 2,9% al 4,8%), Ceuta (del 3,8 al 5%), País Vasco (del 5,2% al 5,3%), Andalucía (4,7% al 4,9%) y La Comunidad valenciana (del 2,6% al 3,6%).

Por otro lado, las comunidades autónomas que tienen menos asma en 2017 que en 2011 serían: Asturias (del 6,9% al 5,7%), Murcia (del 7,1% al 2,9%), Extremadura (del 4,4% al 4,3%), Canarias (del 4% al 3,8%), Cataluña (del 3,7% al 3,5%) y Baleares (del 5% al 4,2%). Mostrando cifras similares estaría Castilla- La Mancha (4,3%).



Colores con tonalidades rosas-rojo con diferentes intensidades mostrando las comunidades autónomas donde ha aumentado más y en mayor cantidad el asma

Colores verdes con diferentes intensidades representando las comunidades autónomas en las que ha habido menos asma

Color gris muestra las comunidades autónomas donde no ha habido cambios en la evolución del

Figura 92. Mapa de evolución de prevalencia de asma en España según los colores de la leyenda.

Resultados

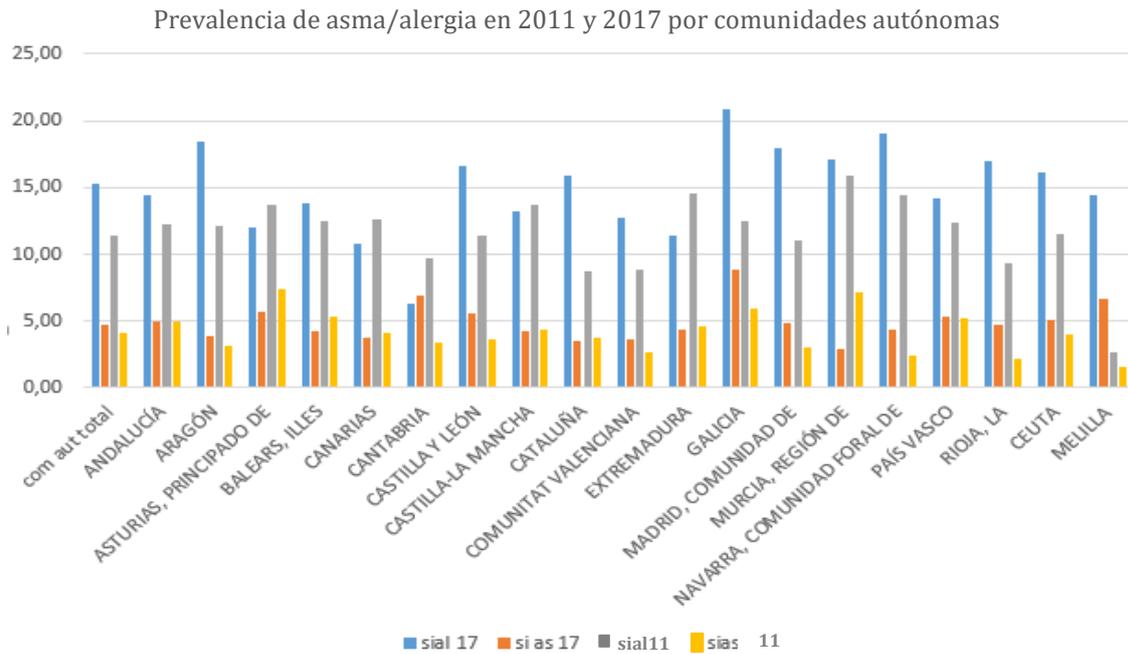


Figura 93. Diagrama de barras mostrando las prevalencias de asma y alergia por comunidades autónomas en los años 2011 y 2017.

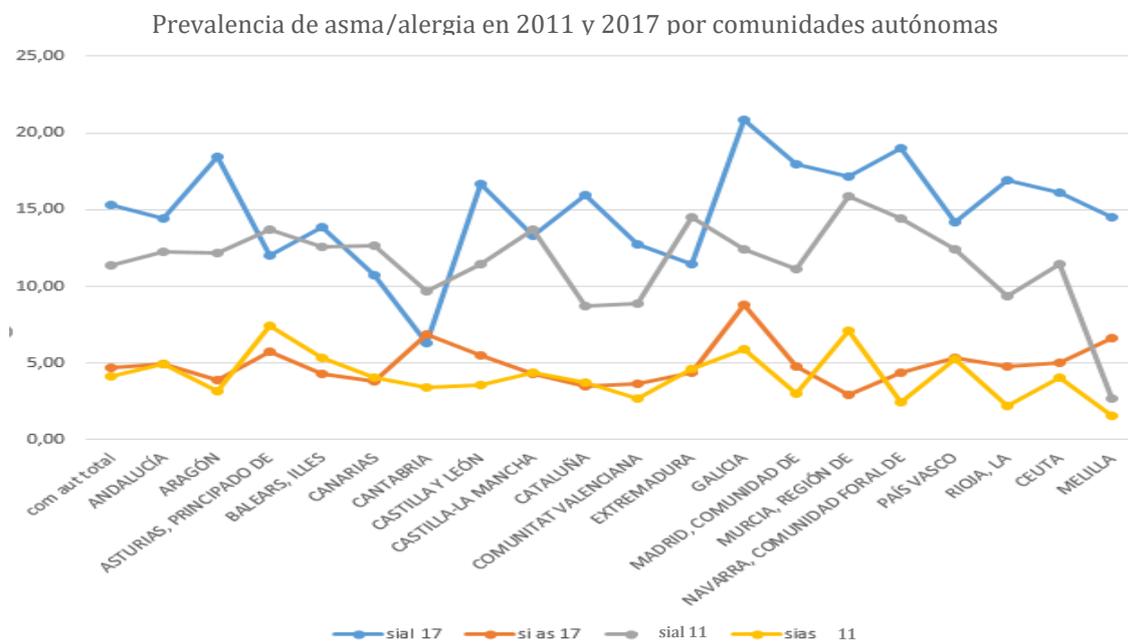


Figura 94. Evolución de la prevalencia de alergia y asma por comunidades autónomas en los años 2011 y 2017.

Las figuras 93 y 94 nos dan una panorámica conjunta, tanto de la alergia como del asma, de la prevalencia y su aumento o disminución en los tiempos de medición citados. Así mismo, nos facilitan visualmente la comparativa de la proporcionalidad de dichas prevalencias.

Podemos ver reflejado que tienen más asma en 2017 que en 2011, sobre todo, en Melilla con un crecimiento de 5,6 puntos (del 1,5% al 6,6%). Con un crecimiento de aproximadamente 3 puntos estaría Cantabria con un 3,5 (del 3,4% al 6,9%) y Galicia con un 3,2 (del 5,6% al 8,8%). Con un crecimiento sobre los dos puntos encontraríamos a La Rioja con un 2,5 (del 2,2% al 4,7%), y Castilla-León con 2 (del 3,5% al 5,5%). Y con aumentos menores de dos puntos: Navarra con un 1,9 (del 2,4 al 4,3%), Madrid con 1,9 (del 2,9% al 4,8%) y Ceuta con un 1,2 (del 3,8 al 5%).

La Comunidad Valenciana tiene ligero aumento de un punto (del 2,6% al 3,6%). Y por último, con pequeños aumentos estaría el País Vasco (del 5,2% al 5,3%) y Andalucía (4,7% a 4,9%).

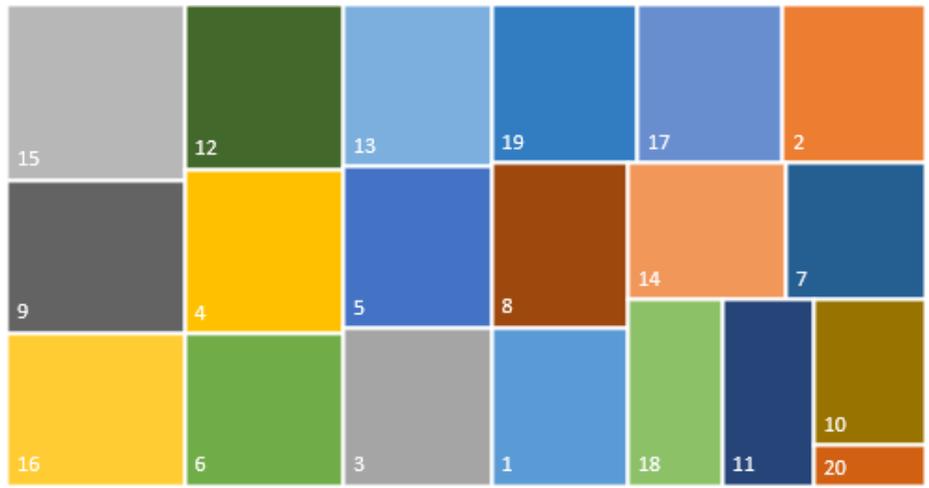
A continuación, en las figuras siguientes podemos visualizar fácilmente por tamaños de los distintos cuadros las diferentes prevalencias de las comunidades autónomas tanto de alergia como de asma en los años 2011 y 2017.



Figura 95. Codificación en mapa de comunidades autónomas.

Resultados

Proporción de prevalencia de alergia por comunidades autónomas en 2011



Proporción de prevalencia de alergia por comunidades autónomas en 2017

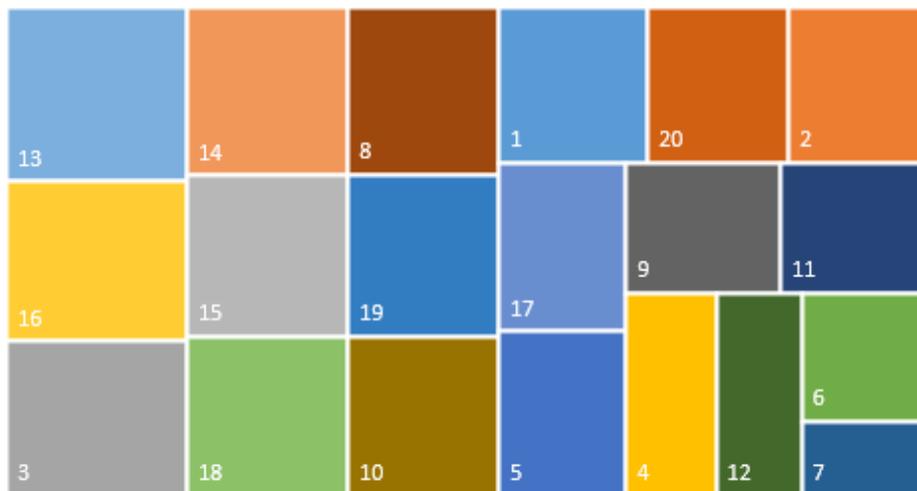
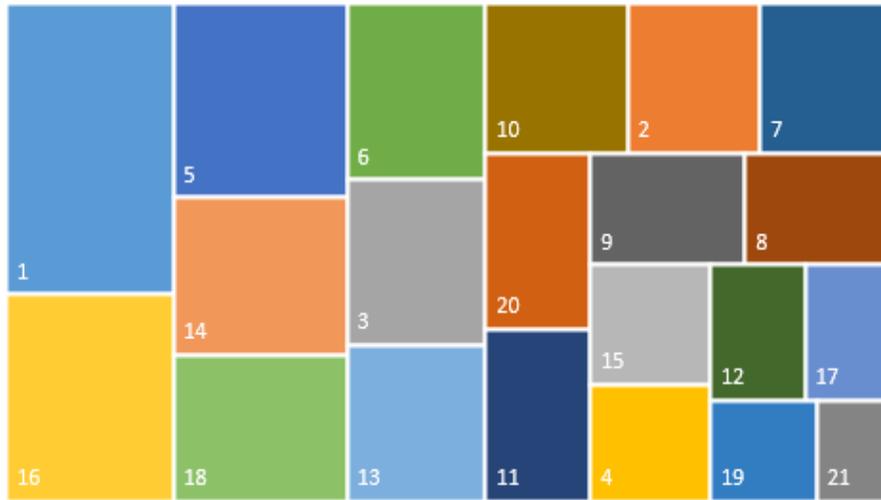


Figura 96. Gráfico de sectores mostrando la proporción de alergia según tamaño de los cuadros por comunidades autónomas en los años 2011 y 2017.

Proporción de prevalencia de asma por comunidades autónomas en 2011



Proporción de prevalencia de asma por comunidades autónomas en 2017

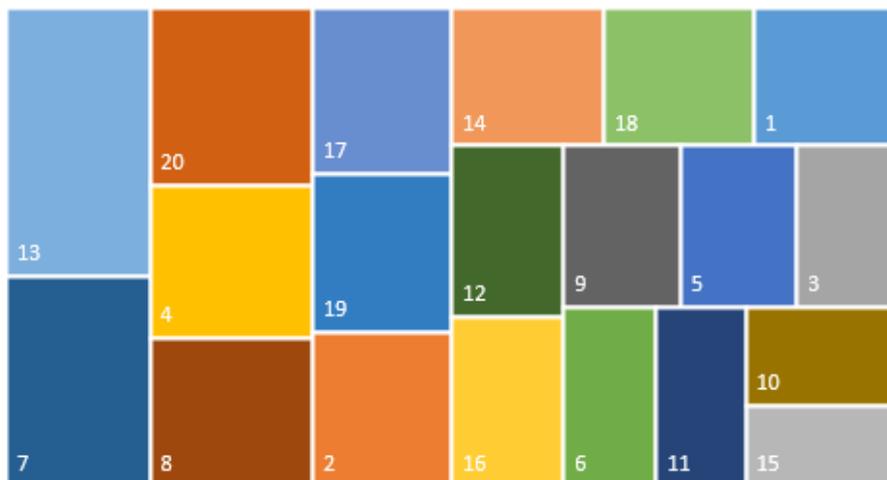


Figura 97. Gráfico de sectores mostrando la proporción de asma según tamaño de los cuadros por comunidades autónomas en los años 2011 y 2017.

V. DISCUSIÓN

Con nuestro trabajo intentamos aportar datos sobre la epidemiología de la alergia y el asma, ya que es difícil encontrar estudios sobre este aspecto en nuestro país. Nuestro primer objetivo era destacar la utilidad de la gran cantidad de datos agregados que existen en general sobre cuestiones médicas y resaltar el escaso aprovechamiento que se hace de ellos por los profesionales y aplicar esta situación al estudio de la alergia y asma. Para ello, pensamos en la opción de realizarlo desde los datos agregados acumulados en las estadísticas oficiales, como los proporcionados por Encuestas Nacionales de Salud y del INE (Instituto Nacional de Estadística).

También estaba dentro del primer objetivo la recopilación y examen de las publicaciones epidemiológicas sobre alergia y asma en nuestro país. Si miramos a los escasos estudios localizados, el antecedente que aporta más datos es Alergológica (156, 157, 158), el cual es un proyecto de la Sociedad Española de Alergología e Inmunología Clínica ante la necesidad de actualizar los datos epidemiológicos de la patología alérgica en España. Pero hay que notar, que el objetivo general de este estudio fue obtener información de la práctica alergológica española sobre los pacientes que consultan por primera vez con un alergólogo. Es decir, no nos aporta cifras reales y generales sobre la epidemiología de la alergia, sino sobre los diagnósticos más frecuentes entre los enfermos que han sido remitidos a consultas de alergia, distribuidas en todo el territorio nacional. Como el propio estudio declara, los objetivos específicos eran analizar las características epidemiológicas y clínicas de la población que acude por primera vez a las consultas de alergología españolas y describir la prevalencia de los distintos diagnósticos (rinitis, asma, urticaria, angioedema, dermatitis, alergia alimentaria, alergia medicamentosa, etc). También analizar las posibles diferencias de distribución regional de la patología alérgica ya diagnosticada y describir el proceder diagnóstico del especialista. El estudio era epidemiológico observacional descriptivo de tipo transversal. Para la selección de los pacientes se realizó una estrategia de muestreo consecutivo de base temporal, durante el periodo de trabajo propuesto a cada investigador se le asignó una fecha aleatoria a partir de la cual comenzó la inclusión

consecutiva de los nuevos pacientes hasta conseguir completar una cifra de 12. En cuanto a resultados, el diagnóstico de rinitis mostró una proporción de 54,1% y el asma de 19,3%. Por lo que respecta a la afluencia de los pacientes que acudían a consulta, a nivel nacional predominaban según nivel de formación los estudios primarios en 37,3% mientras que los estudios medios y superiores eran de un 14,1% y un 9,3%. Los resultados en la Comunidad Valenciana revelaron que los ácaros y los pólenes son los alérgenos más frecuentemente implicados. Al asma se le atribuyó una frecuencia de consulta de un 20,5%. En Valencia, los ácaros representaban en el asma un porcentaje ligeramente superior al 50% y los distintos pólenes una frecuencia que iban de un 5% a un 15%.

En otro de los estudios de epidemiología alérgica realizado en España (159) y que por metodología se asemeja más al efectuado por nuestra parte, se admitía que en los últimos decenios se estaba registrando un aumento de las enfermedades alérgicas en todo el mundo. También insistía en la escasez de estudios de prevalencia de alergia en la población general española. En este trabajo se intentaba establecer la prevalencia de enfermedades alérgicas en una muestra de población general adulta de España, así como establecer las causas y las manifestaciones de esta alergia.

Se realizó un estudio observacional de tipo transversal con un tamaño de muestra que fue de 5.003 sujetos. La muestra se seleccionó de forma aleatoria y automática del listín de teléfono, aplicando el control de cuotas de zona, el número de habitantes, el sexo y la edad. La recogida de datos se efectuó mediante encuesta telefónica siguiendo la técnica CATI (Computer-assisted Telephone Interview technique).

Los resultados mostraron: de los 4.949 sujetos incluidos en el análisis, un 21,6% refirieron ser alérgicos. La prevalencia fue mayor en mujeres (24,6%), en el grupo de 18 a 24 años de edad (26,9%) y en poblaciones de más de 500.000 habitantes (24,3%), mientras que fue menor en los varones (18,3%), en la región norte-centro (17,8%) y en poblaciones con menos de 10.000 habitantes (18,7%). No hubo diferencias significativas en relación al nivel socioeconómico. La rinoconjuntivitis fue la manifestación alérgica más frecuente (45,4%) seguida del asma bronquial (24,9%), la urticaria (24,6%), la dermatitis (21,5%) y el angioedema (6%). Las tres

causas más frecuentes de alergia fueron con diferencia los pólenes (31,5%), los medicamentos (29,4%) y los ácaros del polvo doméstico (25,3%). Mucho menos frecuentes fueron los animales (6,8%), los metales (4,9%), los alimentos (4,8%), los hongos (3%), las picaduras de insectos (2,5%), la exposición al sol (1,9%) y el látex (0,8%). Las conclusiones fueron que la prevalencia de alergia en la población adulta española es del 21,6%, y es más frecuente en las mujeres y en los núcleos urbanos de más de 500.000 habitantes. La manifestación más frecuente es la rinoconjuntivitis. Las causas más frecuentes de alergia son los aeroalérgenos (pólenes y ácaros) y los medicamentos.

En general, a nivel mundial se reconoce que no existen muchos datos recientes sobre la alergia y el asma en adultos. El asma, la rinitis y el eccema son tres afecciones crónicas comunes, si buscamos los datos más recientes a nivel europeo obtenemos que no ha habido datos recientes de varios países sobre la carga de estas tres afecciones en adultos. La fase I de la Red Global de Asma es un estudio poblacional transversal multinacional que utiliza la misma metodología básica que el Estudio Internacional de Asma y Alergias en la fase III Infantil. Proporciona datos sobre la carga de asma, rinitis y eccema en niños, adolescentes y en adultos. Los objetivos del estudio (234) eran llenar este vacío de evidencia. Se disponía de datos de 193.912 adultos (104.061 mujeres; edad media \pm 38 \pm 7,5 años) en 43 centros de 17 países. La heterogeneidad fue alta entre países. La prevalencia de asma como antecedente o con aparición ocasional del 4,4%; oscilando, según países, entre el 0,9% y el 29%. Las mujeres tuvieron una prevalencia más alta de asma que los hombres (4,8% *versus* 3,9%). La prevalencia de la rinitis como antecedente o de aparición ocasional fue del 14,4%, oscilando entre el 2,8% y el 45,7%. Fue mayor en las mujeres que en los varones (14,7% *versus* 14%). En el caso del eccema las cifras fueron de 9,9% (1,6-29,5%).

Hubo una correlación moderada entre la rinitis y el asma, y entre el eccema y la rinitis a nivel del lugar donde se realizaba el muestreo. Hubo correlaciones moderadas a fuertes entre la prevalencia de síntomas de asma, rinitis y eccema detectado en adultos y grupos de edad más jóvenes. La carga de las tres condiciones fue mayor en los países de mayores ingresos.

Se concluyó que existe una carga mundial sustancial de asma, rinitis y eccema en adultos, lo que representa un importante problema de salud pública mundial. Se subrayaba la necesidad de estrategias accesibles, asequibles, equitativas y eficaces para reducir esta carga a lo largo de toda la vida.

Si comparamos con nuestros datos, la prevalencia del asma entre el 4,4-6,6% sería similar a nuestras cifras (4-5%) y el 14,4% para la rinitis alérgica será comparable al 15% de nuestros datos. Si que había similitud en la mayor prevalencia en mujeres y también el crecimiento de la alergia en todos los grupos de población.

Al contrario que en adultos, si que ha habido preocupación por conocer la prevalencia mundial de las alteraciones alérgicas infantiles y su presunto incremento. El Estudio Internacional de Asma y Alergia en la Infancia (ISAAC) (160) investigó la prevalencia de los síntomas y los determinantes del asma, la rinoconjuntivitis y el eccema en escolares en dos ocasiones anteriores (ISAAC fase I en 1993-1995 e ISAAC fase III en 2001-2003). Los objetivos del estudio ISAAC (International Study of Asthma and Allergy in Children) fueron: 1. Describir la prevalencia y gravedad del asma, rinitis y eccema en los adolescentes y compararlas entre países. 2. Obtener mediciones basales para valorar la tendencia futura de esta prevalencia y gravedad. 3. Proporcionar un marco para una posterior investigación etiológica sobre los factores que afectan a estas enfermedades. Los últimos resultados del estudio ISAAC (fase III) muestran que el asma presenta estabilización o descenso y que ha ocurrido principalmente en los países desarrollados, mientras que en los países en desarrollo continúa aumentando. Las causas de estos cambios aún son desconocidas. Una de las explicaciones con más aceptación es que factores ambientales, en especial motivados por la «occidentalización» de la forma de vida, están modificando la respuesta inmunológica y se especula que la etapa prenatal o los primeros años de vida son los periodos más susceptibles para que ocurran estos cambios (179). En cuanto a la rinitis, en ISAAC fase III se observó un ligero aumento mundial en la prevalencia de rinoconjuntivitis, pero las variaciones fueron grandes entre las diversas zonas estudiadas y no hubo un patrón regional consistente (180). En España, según los distintos estudios (182, 183, 184, 238) del International Study

of Asthma and Allergies in Childhood (ISAAC), la prevalencia de rinitis alérgica es del 8,5% entre los seis a siete años y del 16,3% entre los 13 a 14 años, existiendo gran variabilidad entre zonas geográficas, debido a diferencias en los factores ambientales. Esto supone un alto coste económico, sobre todo por las comorbilidades asociadas, especialmente el asma.

El trabajo de ISAAC a nivel mundial ha sido continuado por la Red Global de Asma (GAN) (182, 234). Este es un estudio transversal basado en la población multinacional diseñado para evaluar las tres condiciones y su gravedad, manejo y factores de riesgo en adolescentes de 13 a 14 años, niños de 6 a 7 años y sus padres / tutores utilizando los mismos métodos que ISAAC Fase III. Para aclarar estos aspectos se diseñó una macroencuesta en la que participaron centros pertenecientes a varios países y los resultados globales aparecieron en diferentes publicaciones. También se comunicaron datos parciales correspondientes a centros españoles y extranjeros, a veces ampliados con investigaciones complementarias. La heterogeneidad fue alta entre países. La prevalencia de asma como antecedente o con aparición ocasional del 4,4% y la prevalencia de rinitis fue del 14,4%.

Los cambios en la epidemiología y en la práctica clínica de la alergia en población infantil en España ha merecido otro trabajo (238) que resumía alguno de los datos que tenemos en nuestro país. En las últimas décadas, las enfermedades alérgicas han aumentado de forma exponencial y, aunque parece que la prevalencia del asma en pediatría se está estabilizando (en España se estima en torno al 10%) no ocurre lo mismo con la alergia alimentaria y la anafilaxia que están en claro incremento, constituyendo un problema de salud pública de primera magnitud. En las últimas décadas, se registra un aumento notorio en la alergia alimentaria, más manifiesto en los países industrializados. El pico de prevalencia se presenta al año siendo de 6 a 8%, posteriormente, va disminuyendo hasta el final de la niñez, donde alcanza valores del 3 al 4%, que se mantendrán estables a lo largo de los años. La prevalencia de la alergia alimentaria primaria parece permanecer estable, pero se observa un aumento en la prevalencia de reacciones por reactividad cruzada. Aunque la mayoría de los estudios de prevalencia se basan en resultados sobre sensibilización alimentaria, la prevalencia publicada de alergia puede estar sobreestimada. Los alimentos implicados suelen ser los más consumidos. En cuanto a la

hipersensibilidad a fármacos, un 10% de padres refieren que su hijo/a ha tenido al menos una reacción, constatándose en pocos casos tras un estudio alergológico. Los fármacos implicados con mayor frecuencia son los antibióticos betalactámicos, seguidos de los antiinflamatorios no esteroideos y los antibióticos no betalactámicos. Considerando las tendencias epidemiológicas, las predicciones de la Academia Europa de Alergología e Inmunología Clínica (EAACI) estiman que, en menos de 15 años más de la mitad de la población europea padecerá algún tipo de alergia. El aumento en la patología alérgica puede tener relación con múltiples aspectos como el avance en las técnicas diagnósticas, el descubrimiento de nuevos alérgenos, la mayor concienciación de la población, la contaminación ambiental, modificaciones en los hábitos dietéticos y la teoría higienista, entre otros. Dicho incremento en la prevalencia ha ido en paralelo con un progreso en los métodos diagnósticos y terapéuticos.

En nuestro estudio, en su segundo objetivo declarado, se buscaba determinar la prevalencia de alergia y asma a nivel nacional. No solo eso, también intentaba aclarar, en el apartado primero de este segundo objetivo, si realmente las series temporales nos muestran la idea preconcebida de que alergia y asma están aumentando en los últimos años.

Hemos puesto de manifiesto una serie temporal de prevalencia de alergia desde 1993 a 2017. Lo más destacable es el aumento de prevalencia en 25 años desde un 6,25% a un 15,28%, el cual además es significativo estadísticamente, como muestran los cálculos de la T de Student, tomando como referencia las cifras en 1993 y en 2017. Vemos que el ascenso es progresivo y no se estabiliza en ninguno de los años de medición. De esta forma tan clara se visualiza y cuantifica el lugar común del aumento de la alergia con los años. La idea de que la alergia ha aumentado de forma imparable en los últimos años siempre ha estado ahí y se admitía tácitamente esto, pero visualizamos de forma nítida que tenemos unos datos cuantitativos apoyando este aumento.

Nos planteábamos, en el segundo apartado del segundo objetivo, conocer que grupos de edad y sexo se afectan con más frecuencia. Detectamos en todos los

grupos de edad y sexo que el aumento ha sido constante en el tiempo. Cuando llegamos a 2017, las cifras de prevalencia son en todos los nichos por edad y sexo claramente superiores a los de los años anteriores. El grupo de 15-24 años destaca de forma uniforme en todos los años siendo el nicho al que afecta con más frecuencia la alergia. Esta es una idea común (la alergia es una enfermedad de gente joven) pero aquí aportamos datos cuantitativos de forma irrefutable. Como es lógico en los nichos de edad más avanzados se observa un claro descenso de afectación por alergia. En el caso de los varones el segundo grupo más afectado es el de la década de 25-34 años que además lo hace de forma homogénea en todos los años. En mujeres no es tan evidente la preeminencia del grupo de 15-24 años, siendo superado en 2014 y 2017 por la década de 25-34 años. En general, hay menos diferencias en las distintas décadas siendo la distribución más uniforme. La prevalencia en varones es claramente mayor en el grupo de 15-24 años con respecto a los demás grupos de edad, Desde ese grupo es como una escalera que desciende progresivamente hasta los grupos de mayor edad. En 2017 y en los años anteriores se muestra un dibujo similar excepto en 2014 en que el grupo de 25-34 años era ligeramente superior. La idea común de que la alergia afecta en mayor medida a las mujeres también es habitual pero en los números y gráficos los datos muestran claramente la diferencia a favor de las mujeres en cuanto a prevalencia y esto se da en todos los años.

Si comparamos con el estudio de P. Gaig *et al* (159) en que los resultados mostraron: de los 4.949 sujetos incluidos en el análisis, un 21,6% refirieron ser alérgicos. La prevalencia fue mayor en mujeres (24,6%), en el grupo de 18 a 24 años de edad (26,9%) y en poblaciones de más de 500.000 habitantes (24,3%). Vemos que en nuestro caso la cifra de prevalencia es alta aunque menor que en el indicado estudio, pero si hay semejanzas con el predominio femenino y en los grupos de edad jóvenes, así como una proporción mayor en ciudades muy pobladas.

En el apartado tercero dentro de nuestro segundo objetivo, se trataba de investigar si determinados factores sociales (tipo de población, condición social, económica, laboral, y nivel de estudios) influyen en la prevalencia.

Nuestros datos sugieren un aumento de alergia en estratos económicos más altos junto con una asociación con un mayor nivel educativo. Lo contrario ocurriría en el asma.

Si tenemos en cuenta los niveles de estudio, considerando como tales: sin estudios, primer grado, segundo grado primer ciclo, segundo grado segundo ciclo y tercer grado; con nuestro trabajo podemos ver que los valores de alergia son más altos en personas con nivel de tercer grado, tanto en cifras totales (14,71%) como si separamos en varones (13,70%) y mujeres (15,68%), ocurriendo este hecho independientemente del año en que se analice. Las cifras más bajas de prevalencia coinciden con el grupo de personas sin estudios (10,84% en total, 7,90% en varones, 13,15% en mujeres) y que las más altas son para los niveles superiores de tercer grado. De forma que tanto en cifras totales como para varones y en menos grado en mujeres se forma una escalera ascendente de prevalencia de alergia hacia los niveles más altos de estudios.

Otros estudios a nivel europeo no han usado los niveles educativos para investigar su relación con la alergia o el asma, pero si que han intentado relacionar el nivel educativo de los padres y su influencia en el diagnóstico de asma y alergia en sus hijos. La educación es un buen indicador de posición socioeconómica y se considera capaz de predecir otros indicadores indirectos de posición socioeconómica como ingresos y ocupación. En particular, los bajos niveles de educación materna se han asociado con una mala salud general, obesidad y alto riesgo de enfermedades respiratorias como el asma en los niños. En general, se admite la posición socioeconómica de los padres como determinante conocido de la salud de un niño, y como sustituto de la posición socioeconómica se ha considerado la educación parental.

Una publicación (235), con un análisis secundario de datos de dos estudios transversales realizados en Italia en 2006, tenía como objetivo investigar si una educación parental baja producía un efecto directo sobre el asma y la rinitis alérgica en los niños, o si las asociaciones están mediadas por la exposición a otros factores de riesgo personales o ambientales. Los datos de 2.687 adolescentes (10-14 años) se analizaron mediante un modelo de análisis de trayectoria utilizando modelos de

ecuaciones estructurales generalizadas. A través del análisis de la trayectoria, se mostró un efecto indirecto de la educación de los padres sobre el asma en niños, mediado por el tabaquismo materno durante el embarazo, y sobre la rinitis mediada por un entorno tabáquico temprano. También se encontró que la educación de los padres no afectó directamente el riesgo de que sus hijos recibieran un diagnóstico de asma ni tuvieran rinitis, y que aunque influyó significativamente en el número de niños en la familia y el índice de hacinamiento, los efectos indirectos mediados por estos factores de riesgo no fueron significativos en esos resultados. No se encontraron efectos directos o indirectos significativos de la educación de los padres en el asma. Los autores reconocen que ningún estudio ha extrapolado el efecto indirecto de la educación de los padres, sobre el asma. Estos resultados sugieren la importancia de la promoción de políticas de salud *ad-hoc* para promover el abandono del hábito de fumar, especialmente durante el embarazo.

Los datos sobre la relación entre posición socioeconómica y las enfermedades alérgicas respiratorias en los niños aún son inconsistentes y controvertidos. Por un lado, existe evidencia de que la alta posición socioeconómica de los padres se asocia con una mejor salud infantil, aunque la sensibilización a los alérgenos transportados por el aire parece ser más frecuente en individuos con un nivel socioeconómico más alto en la infancia. Por otro lado, los grupos socioeconómicamente desfavorecidos generalmente enfrentan una mayor exposición a factores de riesgo ambientales, lo que puede aumentar la susceptibilidad al desarrollo de enfermedades respiratorias alérgicas y el asma infantil sería más común en familias con bajo nivel socioeconómico. Se puede argumentar que los padres con posición socioeconómica más baja a menudo adoptan comportamientos poco saludables como fumar durante el embarazo y continuar haciéndolo después.

Alguna publicación (236) ha analizado, en forma de cohorte, la relación entre el estado socioeconómico y el asma junto con la rinitis. Así, cuatro mil ochenta y nueve familias con hijos nacidos entre 1994 y 1996 en áreas predefinidas de Estocolmo respondieron cuestionarios sobre factores ambientales, estado socioeconómico (ocupación parental) y síntomas de enfermedad alérgica al nacer, en determinados momentos de edad de los niños (1, 2 y 4 años de edad). Las muestras de sangre tomadas a los 4 años de 2.614 niños se analizaron para detectar IgE específica para

alérgenos comunes en el aire y alimentos. Hubo una disminución del riesgo de asma y rinitis con el aumento del nivel socioeconómico. El riesgo de sensibilización a los alérgenos alimentarios también disminuyó con el aumento del nivel socioeconómico. Se concluyó que el asma y la rinitis son más comunes en los grupos socioeconómicos más bajos después del ajuste por factores de riesgo tradicionales. Esto podría estar relacionado con diferencias adicionales no controladas en el estilo de vida y las exposiciones ambientales entre los grupos, y requeriría de más evaluaciones.

Por lo referido, la bibliografía mencionada no coincide totalmente con nuestros resultados. Estos muestran niveles superiores de alergia, de unos dos puntos, en los niveles de renta más altos con respecto a los niveles de renta más bajos con significación estadística. Esto es independiente del año en que se realiza la medición. La capacidad económica más alta hace que la alergia suba dos puntos, más claramente en varones. Es más marcado en la población entre 15-24 años y estos datos se acentúan en los últimos años.

Si que coincide con nuestra aportación, alguna revisión sistemática de trabajos publicados acerca de la posición socioeconómica en el desarrollo del alergia. En una de ellas (237), de los 4.407 registros bibliográficos identificados, 183 se incluyeron en el análisis. Las investigaciones sobre alergia mostraron una asociación positiva entre un mayor nivel económico y la enfermedad en el 66% de los estudios.

En nuestro segundo objetivo, para realizar el análisis estadístico de la prevalencia de asma (dentro del primer apartado), hemos escogido cuatro fechas de medición: 2006, 2011, 2014, 2017, debido a que en ese momento es cuando el INE separa los datos de asma de los datos de bronquitis crónica y enfisema, que eran recogidos conjuntamente hasta ese momento. La prevalencia del asma no ha sufrido la misma evolución que la prevalencia de alergia. Como vemos en los sucesivos desarrollos las cifras de asma se han mantenido estables en los últimos 12 años. Teniendo en cuenta solo la edad, el asma de 2006 a 2017 pasa de 4,45% a 4,68%.

Si pasamos al segundo apartado, la prevalencia por grupos de edad y sexo en el asma, en la población de más de 75 años es de 7,77% en 2006 y de 6,51% en 2017. En dicha población en varones llega a ser de 8,94% en 2006. En mujeres es menos acusado pero también hay un máximo de 7% en 2006. En la población de más de 75 años es de 7,77% en 2006 y de 6,51% en 2017. En varones se observan aumentos de hasta 9% en 2006 en el grupo de 65-74 años. En las poblaciones de más de 75 años se alcanzan cifras de 6% que contrasta con las cifras en varones en las poblaciones de 25-44 y de 55-64 años que son del 2%.

Llama la atención la mayor prevalencia en las poblaciones de más edad, y por otro lado la uniformidad en los puntos de medición a lo largo de los 12 años. El aumento en la población de más de 75 años tanto en cifras totales como particularmente en varones y mujeres es evidente, así como la caída en grupos de 45-54 años. Este dibujo nos impacta porque siempre se ha dicho que el asma es una enfermedad crónica por lo que se debería mantener en una subida progresiva en el tiempo y no con baches. La serie temporal sobre el asma señala no un ascenso lento sino una estabilización de las cifras con respecto a 12 años atrás, con una caída en edades medias dibujando una "V".

Las referencias bibliográficas admiten que el asma es un trastorno crónico común que puede estar aumentando en prevalencia. Sin embargo, poco se sabe de su distribución y determinantes. La Encuesta de Salud Respiratoria de la Comunidad Europea (234) es una encuesta multicéntrica sobre la prevalencia, los determinantes y el tratamiento del asma. Presenta una descripción de la variación de los síntomas del asma en toda Europa. Se distribuyó un cuestionario de detección, que incluía siete preguntas relacionadas con la prevalencia de 12 meses de síntomas de asma, a muestras representativas de hombres y mujeres de 20 a 44 años de edad en 48 centros, predominantemente en Europa occidental. La media de la tasa de respuesta al cuestionario fue del 75%. La prevalencia de todos los síntomas varió ampliamente. Aunque estos fueron generalmente más bajos en el norte, centro y sur de Europa y más altos en las Islas Británicas, hubo grandes variaciones incluso dentro de algunos países. Los centros con una alta prevalencia de ataques de asma también informaron altas prevalencias de alergias nasales. El uso de medicamentos para el asma fue más común donde las sibilancias y los ataques de asma fueron más

frecuentes. En la mayoría de los centros de los Países Bajos, Suecia y el Reino Unido, más del 80% de las personas con un diagnóstico de asma estaban usando actualmente medicamentos para el asma. En Italia, Francia y España, la tasa fue generalmente inferior al 70%. Estos datos son la mejor evidencia hasta la fecha de que existen diferencias geográficas en la prevalencia del asma, son sustanciales y no son un artefacto del uso de métodos no comparables.

Las preguntas utilizadas en él fueron las siguientes: "¿Ha tenido sibilancias o silbidos en el pecho en algún momento en los últimos 12 meses?" y "¿Ha tenido un ataque de asma en los últimos 12 meses?". Hubo grandes variaciones geográficas en la prevalencia del asma. La media de prevalencia del asma fue del 4,5% (rango 2-11,9%) en la etapa uno y del 5,2% (rango 1,2-13%) en la etapa dos. Las mujeres tenían una mayor prevalencia de asma que los varones y la prevalencia de sibilancias se asoció negativamente con la edad. Se concluyó que las variaciones geográficas en la prevalencia del asma probablemente se debieron a factores ambientales.

Llegando al tercer apartado de nuestro segundo objetivo, si comparamos con los resultados obtenidos en otros trabajos, encontramos que no existen referencias en la bibliografía de epidemiología del asma en relación con factores tales como tipo de población o factores laborales como tipo de ocupación y tipo de jornada laboral, aunque si alguno que trata de forma muy generalista el asma desde perspectivas socioeconómicas.

Algunos estudios (236) han propuesto que el asma infantil es más común en familias con bajo nivel socioeconómico. En el estudio ya mencionado realizado en Suecia se evaluó la relación entre el estado socioeconómico y el asma y la rinitis. Se utilizaron cuestionarios sobre factores ambientales, estado socioeconómico (ocupación parental) y síntomas de enfermedad alérgica al nacer (cuando los niños tenían 1, 2 y 4 años de edad). Las muestras de sangre tomadas a los 4 años de 2.614 niños se analizaron para detectar IgE específica para aeroalérgenos comunes y alimentarios. Se detectó que el asma es más común en los grupos socioeconómicos más bajos que en los más altos.

También se han realizado revisiones sistemáticas de trabajos publicados acerca de la posición socioeconómica en el desarrollo del asma. En una de ellas (237), de los 4.407 registros bibliográficos identificados, 183 se incluyeron en el análisis. Una clase social baja se asoció con una mayor prevalencia de asma en el 63% de los estudios.

En nuestro trabajo, las conclusiones son en parte semejantes, los niveles económicos más altos muestran bajos valores de asma (como nos dicen las prevalencias encontradas 2,11% en 2006, 2,93% en 2011, 3,46% en 2014, 4% en 2017).

En nuestra aportación, hemos relacionado el nivel educativo de las personas afectas de asma con su patología. En cifras totales de ambos sexos el grupo sin estudios representa el 6,76% y el grupo de nivel educativo más alto el 3,84%, es decir la cifra más alta y la más baja respectivamente. En mujeres la diferencia es mayor, un 7,17% sin estudios y un 4,15% en el nivel más alto.

Si tenemos en cuenta el nivel de estudios (con la clasificación ya mencionada) vemos que en cifras totales el grupo de sin estudios tiene prevalencias más altas: 7,80% en 2006, 6,62% en 2011, 7,33% en 2017. Sobre todo con respecto al grupo de nivel más alto que oscila entre 3,33% en 2006, 3,55% en 2011, 3,88% en 2014 y 4,63% en 2017.

Los resultados son claros, en varones sin estudios las cifras son más altas, de casi el 9% en 2006. En mujeres sin estudios el aumento es de un 30% más. Hay un marcado aumento en la población sin estudios, tanto de ambos sexos como por separado.

Una conclusión diferente se encuentra en otros trabajos (235), los cuales reconocen al nivel educativo de los padres como un sustituto de posición socioeconómica y lo utilizan como un indicador. Se encontró que la educación de los padres no afectó directamente el riesgo de que los niños recibieran un diagnóstico de asma. Sin embargo, revelamos un efecto indirecto en el número de niños en la familia y el índice de hacinamiento. Los padres con posición socioeconómica más baja a menudo adoptan comportamientos poco saludables como fumar durante el embarazo y continuar haciéndolo después. Por otro lado, el estilo de vida favorable de los padres con posición socioeconómica alta parece prevenir el asma.

No hemos encontrado datos de trabajos que relacionen, como si lo hemos efectuado nosotros, el asma con el número de habitantes del núcleo de población en que se reside. No hemos detectado diferencias significativas en el total de ambos sexos en asmáticos según población. Las cifras más bajas a nivel total (ambos sexos) corresponden a poblaciones de más de 400.000 habitantes (3,95%) mientras que en poblaciones de hasta 10.000 habitantes es de 4,48%. Por separado se repite el patrón.

En poblaciones de más de 400.000 habitantes se dan las cifras más bajas a nivel total y también en mujeres. Las cifras más altas son en poblaciones de 100.000 a 400.000 habitantes, aunque en el caso de varones las cifras más altas corresponden a poblaciones de hasta 10.000 habitantes. En mujeres la cifra más alta, en todos los años, es en poblaciones de 100.000 a 400.000 habitantes.

Con respecto a nuestro tercer objetivo que era saber si las zonas más contaminadas, como son las grandes ciudades muestran un mayor nivel de alergia y asma, hay que mencionar que un punto importante es la idea generalizada de que la mayor contaminación y el calentamiento global producen un aumento de la alergia y el asma (239). En los últimos 50 años, la temperatura global de la tierra ha aumentado notablemente debido a la creciente emisión de concentraciones de gases de efecto invernadero generados por la actividad humana. Los principales cambios que afectan a la atmósfera y al clima, incluido el calentamiento global, tienen un gran impacto en la biosfera y en nuestro entorno. El incremento de las áreas residenciales y los altos niveles de emisiones vehiculares inducirían síntomas de enfermedades que cursan con obstrucción bronquial (en particular asma bronquial), con mayor prevalencia en personas que viven en áreas urbanas en comparación con aquellos que viven en áreas rurales (240). El aumento de la producción y concentración de aeroalérgenos y contaminantes atmosféricos y los cambios en su distribución geográfica y los cambios en su distribución geográfica, entran en las consecuencias observadas y previstas (241).

Se da por supuesto de que la creciente frecuencia de enfermedades respiratorias obstructivas, en particular del asma alérgica, observada en los últimos años, puede

explicarse parcialmente por diversos cambios ocurridos en el medio ambiente, con presencia creciente en la atmósfera de sustancias químicas (material particulado y componentes gaseosos como dióxido de nitrógeno y ozono) y por factores desencadenantes biológicos (aeroalérgenos). Varios trabajos hacen hincapié en este aspecto (240, 241). Pero hay que remarcar que todas estas publicaciones no distinguen entre patología mediada por mecanismos alérgicos y patología obstructiva o debida a otros factores (humos, irritantes químicos, etc). No podemos dejar de recordar que en la reunificación de Alemania, un hallazgo relevante fue el de una menor prevalencia de respuestas cutáneas positivas en las zonas contaminadas de las ciudades de Alemania del Este (23, 24). Parecía haber un exceso de síntomas por irritantes químicos (incluyendo bronquitis) en las áreas más intensamente contaminadas de la República Democrática Alemana (RDA) equilibrado con una menor prevalencia de asma alérgica (25). También ha habido una preocupación respecto a que la exposición al tráfico denso que podría aumentar el riesgo de asma (52).

A pesar de que los datos son contradictorios, se tiene la idea preconcebida que el cambio climático modifica la producción y dispersión de componentes químicos y biológicos de la contaminación atmosférica. En particular, aumenta el contenido alérgico atmosférico de polen y las esporas, con efectos proinflamatorios en las vías respiratorias de los pacientes alérgicos. También, se teoriza a nivel general que los aeroalérgenos modificados por la contaminación serían más capaces de estimular, en sujetos alérgicos, la reactividad de las vías respiratorias, induciendo síntomas de asma bronquial. El papel de los factores ambientales en relación con el asma y la alergia fue cobrando interés a lo largo de la década de 1990 (33). Ha habido una preocupación general creciente respecto a que los cambios en la contaminación ambiental hayan sido la base del incremento del asma, pero los profesionales no están tan seguros de ello (35). La cuestión más importante es si la exposición a largo plazo a la contaminación atmosférica influye en la prevalencia del asma. Esto ha sido evaluado en estudios por áreas en los que se ha comparado la prevalencia de asma en toda una ciudad o en ciudades con mezclas diferentes de contaminación (34, 35).

Se ha intentado relacionar el efecto de las partículas existentes en atmósferas contaminadas con la aparición de alergia y asma (50). Los contaminantes que han

sido determinados con mayor frecuencia en las comparaciones geográficas para relacionar la prevalencia de la enfermedad con sibilancias y la contaminación atmosférica con el dióxido de azufre y las partículas aéreas (54, 55). Pocos estudios han evaluado los posibles peligros debidos a contaminantes como óxidos de nitrógeno (NO₂) y ozono (58). En general, los valores de NO₂ se correlacionan con las concentraciones de partículas de dióxido de azufre (SO₂) y resulta difícil diferenciar sus efectos independientes (59, 60).

Se postula que el aumento de CO₂ y las partículas diesel producirían un aumento de la agresividad de los pólenes y por tanto un aumento de la alergia (52, 69, 71). En nuestro caso si que se aprecia un aumento de alergia en las poblaciones de más de 400.000 habitantes donde se presupone una mayor contaminación pero si comparamos con el aumento global de alergia, independientemente del tipo de población, que ha sido superior a un 50% podríamos pensar que podría deberse a otros factores, como que los grupos de edad más jóvenes, y que tienen una mayor proporción de alergia, son más abundantes en las grandes ciudades. En las poblaciones de más de 400.000 habitantes el crecimiento de la alergia en los últimos registros ha subido claramente, siendo en mujeres más marcado. En poblaciones de más de 400.000 habitantes de media en 25 años es el 11,52% y ha ido subiendo progresivamente en 25 años. En el sector de 15-24 años se sigue también este patrón. Por poblaciones, vemos que en los últimos años ha aumentado sobre todo la prevalencia en ciudades de 100.000 a 400.000 habitantes y en poblaciones de más de 400.000 habitantes. Mientras que en el resto de poblaciones más pequeñas las cifras son más similares y con menos diferencias entre ellas. Esto es más acusado en el grupo femenino. A través de las series de años varía el porcentaje de prevalencia, viendo que las poblaciones de más de 400.000 habitantes están en los valores más altos mientras que en la década de 2003-2011 se mantenía constante independientemente del número de habitantes. A continuación, vendrían las de 100.000 a 400.000 (10,49%), ocupando el lugar más bajo las que no alcanzan 10.000 habitantes (8,90%) y de 10.000-50.000 (9,52%). No encontramos registros similares en la literatura de asma y alergia de estudios por grupos de población, a lo sumo algunos trabajos que diferencian entre el medio rural y el urbano.

El asma es una enfermedad heterogénea en la que fuertemente influyen factores medioambientales como meteorológicos y el cambio climático, que varían en tipo e intensidad en todo el mundo. Se afirma que la exposición al cambio climático y a la contaminación atmosférica está vinculada a muchos signos de enfermedades respiratorias obstructivas (por ejemplo, exacerbación del asma), aumento del uso de medicación, visitas al servicio de urgencias e ingresos hospitalarios (239).

Nuestros resultados en el caso del asma, no detectan diferencias significativas en el total de ambos sexos en asmáticos según población. Además, las cifras más bajas a nivel total (ambos sexos) corresponden a poblaciones de más de 400.000 habitantes (3,95%) mientras que en poblaciones de hasta 10.000 habitantes es de 4,48%. Por separado se repite el patrón. En poblaciones de más de 400.000 habitantes se dan las cifras más bajas a nivel total y también en mujeres. Las cifras más altas son en poblaciones de 100.000 a 400.000 habitantes, aunque en el caso de varones las cifras más altas corresponden a poblaciones de hasta 10.000 habitantes. En mujeres la cifra más alta, en todos los años, es en poblaciones de 100.000 a 400.000 habitantes.

Nuestro cuarto objetivo era efectuar un estudio ecológico por comunidades autónomas de posibles relaciones de alergia y asma con otros factores de salud como: sedentarismo, actividad física en el trabajo, hábito tabáquico, peso corporal, tipo de desayuno.

La metodología estadística efectuada nos han permitido establecer interesantes relaciones con otros tipos de factores de salud, como el tipo de trabajo en relación con el esfuerzo físico, el sedentarismo, el hábito tabáquico, el normopeso/sobrepeso, etc.

El sedentarismo y la actividad física moderada tendrían una mayor afinidad por la alergia/asma, así como el trabajo sedentario.

En contra, tendrían sobre todo el no sedentarismo en ocio acompañado de una actividad física ligera y el trabajo que requiere gran esfuerzo.

Por tanto, en 2017 la alergia se relaciona con trabajo sedentario y el sedentarismo en ocio. Relación inversa con no sedentarismo y trabajo realizado con gran esfuerzo.

El asma con trabajo de pie y actividad física moderada. Relación inversa con trabajo realizado con gran esfuerzo y trabajo sedentario.

En 2011, la alergia se relaciona con actividad física moderada y trabajo sedentario. Relación inversa con no sedentarismo y actividad física ligera; mientras que el asma se relaciona con actividad física moderada y trabajo con esfuerzo, teniendo relación inversa con no sedentarismo en ocio y actividad física ligera.

En cuanto a otros factores de salud detectamos correlación de asma con trabajo de pie y con población que nunca ha fumado, y correlación inversa con fumador diario. A su vez, alergia tiene correlación con trabajo sedentario, con exfumador y con sobrepeso; correlación inversa con fumador diario.

Nuestro quinto objetivo era determinar la prevalencia de alergia y asma pero diseccionando también datos por comunidades autónomas, y en especial los referidos a la Comunidad Valenciana.

Con respecto a la alergia, en 2011 era Murcia con un 15% quien lideraba la prevalencia de alergia. Entre el 10-15% estaban la mayoría de comunidades: Andalucía, Aragón, Cantabria, Baleares, Canarias, Castilla-León, Castilla-La Mancha, Extremadura, Galicia, Madrid, Navarra, País Vasco y Ceuta. Y con menos del 10% estaban Cantabria, La Rioja, Cataluña y la Comunidad Valenciana. A destacar la baja cifra de Melilla con un 2,2%.

En 2017, Galicia era la comunidad con la prevalencia de alergia más importante (20%), Navarra con un 19% y Aragón con un 18% le seguían. Después aparecían una serie de comunidades con cifras superiores al 15% como Cataluña, Madrid, Castilla-León, Ceuta y Murcia. Con cifras inferiores al 15% una buena parte de las comunidades: País Vasco, Castilla-La Mancha, Extremadura, Canarias, Baleares, Asturias, Andalucía y Melilla. También la Comunidad Valenciana estaría en este grupo con un 11%. A destacar la baja prevalencia en Cantabria con un 6,3%.

Por tanto, la mayoría de comunidades autónomas han aumentado en prevalencia de alergia: Galicia, País Vasco, La Rioja, Navarra, Madrid, Castilla-Leon, Ceuta, Melilla,

Aragón, Baleares, Cataluña, Comunidad Valenciana, Murcia y Andalucía. Todas tienen más alergia en 2017 que en 2011.

Pero el aumento no es el mismo en todos los casos. Es muy llamativo el crecimiento en Melilla, del 2,2% al 14% y de Galicia de 11% al 20%. De la misma forma, con una elevación importante: Aragón que ha pasado de un 11% a un 18% y Navarra de un 13% a un 19%, Cataluña de un 8% a un 15%, Madrid de un 10% a un 17%, La Rioja de un 8,7% a un 16%, Ceuta de un 11% a un 16% y Castilla-León de un 11% a un 16%. Por otro lado, con un crecimiento más moderado: Andalucía del 11% al 14%, País Vasco del 11% al 14%, Baleares de 11% al 13% y Murcia de 15% al 17%. La Comunidad Valenciana ha pasado de un 8,4% a un 12%.

Las comunidades autónomas donde se observa menos alergia en 2017 que en 2011 serían Extremadura (del 12% al 11%), Cantabria (del 9,4% al 6,3%) y Canarias (del 11% al 10%). Asturias se mantiene en un 12% y Castilla-La Mancha se ha mantenido en un 13%.

En lo referente a la prevalencia del asma, tienen más asma en 2017 que en 2011 las siguientes comunidades: sobre todo, Melilla con un crecimiento de 5,6 puntos (del 1,5% al 6,6%) y con un crecimiento de aproximadamente 3 puntos estaría Cantabria con un 3,5 (del 3,4% al 6,9%) y Galicia con un 3,2 (del 5,6% al 8,8%). Con un crecimiento sobre los dos puntos encontraríamos a La Rioja con un 2,5 (del 2,2% al 4,7%) y Castilla-León con 2 (del 3,5% al 5,5%). Y con aumentos menores de dos puntos: Navarra con un 1,9 (del 2,4 al 4,3%), Madrid con 1,9 (del 2,9% al 4,8%) y Ceuta con un 1,2 (del 3,8 al 5%). La Comunidad Valenciana tiene ligero aumento de un punto (del 2,6% al 3,6%). Y por último, con pequeños aumentos estaría el País Vasco (del 5,2% al 5,3%) y Andalucía (4,7% al 4,9%).

Acerca de las comunidades que tienen menos asma en 2017 que en 2011, destacan por su disminución: Asturias (del 6,9% al 5,7%), Murcia (del 7,1% al 2,9%), Extremadura (del 4,4% al 4,3%), Canarias (del 4% al 3,8%), Cataluña (del 3,7% al 3,5%) y Baleares (del 5% al 4,2%). Mostrando cifras similares estaría Castilla-La Mancha (4,3%).

En resumen, la Comunidad Valenciana mostraría, por tanto, un moderado aumento tanto en lo que respecta a la alergia como al asma. La disminución de prevalencia se ha consolidado en Extremadura y Canarias tanto en alergia como en asma. Asturias no varía en alergia y Castilla-La Mancha no varía en asma ni en alergia.

A destacar en sentido negativo, el aumento en Melilla tanto de alergia como de asma y también la consolidación de Galicia en el aumento de ambas patologías.

La cornisa mediterránea comprendiendo Cataluña y Murcia (sin la Comunidad Valenciana) más Baleares disminuirían la prevalencia de asma.

Para conseguir nuestro sexto objetivo que era aplicar técnicas estadísticas como son los estudios multivariantes, y sistematizar el análisis de datos para conseguir resultados valorables (242), hemos introducido esta métodos para el estudio de la epidemiología de la alergia y asma. Las usadas han sido el análisis de componentes principales y el análisis de conglomerados. El análisis multivariante en su versión estadística y su aplicación gráfica permite una visión diferente de múltiples datos clínicos para poder sacar consecuencias prácticas.

En alergia, el análisis de conglomerados ha conseguido cuatro grupos, uno de ellos representaba las prevalencias más altas, otro las prevalencias más bajas y los dos restantes habían integrado las prevalencias medias. Es decir, el análisis de conglomerados nos permitiría visualizar con más rapidez estas diferencias entre grupos y destacar el grupo a buscar que sería donde estaban los valores más elevados de alergia.

En el caso de conglomerados por tipo de población, detectamos que uno de los conglomerados incluía a los nichos de mayor prevalencia, otro a los de menor prevalencia, pero curiosamente en otro de ellos se agrupaban todos los nichos de habitantes de ciudades de más de 400.000 habitantes, independientemente de que pertenecieran a una edad o sexo concreto, que es justamente donde aparecen los mayores valores de alergia.

Los conglomerados obtenidos teniendo en cuenta el nivel educativo, incluían en un caso a los nichos de varones sin estudios y a la poblaciones de mayor edad 45-64

años y más de 65 (sobre todo varones), siendo todos ellos los que muestran las prevalencias más bajas. Es decir, coincidiría con lo reflejado en la estadística descriptiva, la cual es clara, en varones sin estudios las cifras son más bajas. En otra de las agrupaciones estaban las mujeres de grupos de edad de 15-24 y de 25-44 años y también a los varones más jóvenes y con mayor nivel educativo coincidiendo con la estadística descriptiva la cual nos presenta que los valores de alergia son más altos en personas con nivel de tercer grado tanto en cifras totales como si separamos en hombres y mujeres, ocurriendo este hecho independientemente del año en que se analice.

En cuanto a los conglomerados obtenidos en el caso del asma, uno de ellos integra a los nichos de mayor edad y a las mujeres de 15-24 y de 25-34 años, que son precisamente los que presentan una mayor prevalencia. El otro incluye a los varones de 45-54 y 55-64 años que son los de menor prevalencia, junto con el grupo de 45-54 años de ambos sexos que dentro de este apartado es el de menor prevalencia.

La estadística descriptiva nos muestra que en poblaciones de más de 400.000 habitantes se dan las cifras más bajas en totales de ambos sexos y también en mujeres. Las cifras más altas son en poblaciones de 100.000 a 400.000 habitantes aunque en el caso de varones las cifras más altas corresponden a poblaciones de hasta 10.000 habitantes. Sin embargo, estos datos no son relevantes en la composición de los conglomerados, los cuales se distribuyen sin tener en cuenta el tipo de población.

Al realizar el análisis de conglomerados según el nivel de estudios en el paciente asmático, uno de los casos coincidiría con lo reflejado en la estadística descriptiva, habiendo un marcado aumento en la población sin estudios, tanto de ambos sexos como por separado. En este grupo estarían los nichos totales de sin estudios, tanto de ambos sexos como por separado, como diversos nichos separados por edades pero teniendo en común la falta de estudios o la pertenencia a estudios primarios.

Atendiendo al análisis de componentes principales, en su versión gráfica del círculo de correlaciones, los datos en alergia no son relevantes para obtener una rápida visualización del tipo de grupos y de las consecuencias que podemos extraer. Esto es así porque las variables se correlacionan positivamente y existe poco ángulo

entre ellas. Además, la longitud de los vectores en su mayoría es similar por lo que las variables muestran poca diferencia entre ellas. Los grupos que corresponderían al nicho de edad de 15-24 años y por tanto los de mayor prevalencia se sitúan muy próximos en el cuadrante inferior.

Los grupos de edad más alta como el de 75 años ocupan los extremos superior e inferior de los cuadrantes, estando con un ángulo que no llega a ser ortogonal pero es el que más se acerca de toda la distribución y esto nos hablaría de una escasa relación. El grupo de 75 años de varones y el de la misma edad en mujeres, ya que ambos son los de menor prevalencia entre su grupo de género, nos sería indicado por una menor longitud del trazo de la flecha de la variable en el gráfico de influencias.

Si tenemos en cuenta el nivel de población, las variables se agrupan en el eje horizontal del principal componente, no mostrando las variables correlación opuesta excepto algún caso.

En el caso del análisis teniendo en cuenta el nivel de estudios, la mayoría de variables se confunden en su trazado al haber poco ángulo entre ellas lo cual quiere decir que se correlacionan positivamente. Únicamente los nichos de ambos sexos sin estudios de 15-24 años, mujeres de 15-24 años sin estudios y mujeres de 15-24 años con nivel de primer grado, muestran un ángulo recto con respecto al resto de variables expresando su nula correlación. Se trataría de nichos con una prevalencia similar respecto al resto de grupos con nivel de estudios similar, por lo que no es el factor determinante.

En lo referente al análisis de componentes principales en el asma, en el círculo de correlaciones es evidente la diferencia de distribución que veíamos con respecto a la alergia. La longitud de los vectores no es similar por lo que las variables muestran diferencia entre ellas. Las variables están poco correlacionadas porque existe un mayor ángulo entre ellas que el caso de la alergia. Se distribuyen repartidas entre los dos componentes principales que explican la mayor parte de la varianza, con un pequeño aumento en el primer componente principal con respecto al segundo. Los nichos que representan el grupo de edad de 35-44 años en todas sus versiones (ambos sexos, mujeres, varones) se sitúan todos ellos en la zona negativa del

componente principal. En el lado contrario, se sitúan todos los grupos de edad de más de 75 años con una longitud de los vectores importante. Es decir, los menos prevalentes frente a los más prevalentes.

Tras aplicar, según el tipo de población, el análisis de componentes principales en su versión gráfica del círculo de correlaciones, los datos no son relevantes para obtener una rápida visualización del tipo de grupos y de las consecuencias que podemos extraer. Las variables están poco correlacionadas y se distribuyen en abanico porque existe un mayor ángulo entre ellas que el caso de la alergia. La longitud de los vectores tampoco es similar por lo que las variables muestran diferencias entre ellas.

Esta misma situación se produce si aplicamos el nivel de estudios en el asma al análisis de componentes principales.

VI. CONCLUSIONES

Los resultados de la presente investigación permiten llegar a las siguientes conclusiones contestando a los objetivos formulados:

1. La alergia ha aumentado de forma imparable en los últimos años y con nuestro trabajo tenemos unos datos cuantitativos apoyando este aumento. No existían datos pormenorizados con series temporales de 25 años en la bibliografía médica, con respecto a alergia y asma en España.

2. Si valoramos el incremento de prevalencia de la alergia en el periodo considerado en España, observamos que es muy importante, y supera ampliamente más del doble del dato inicial, pasando de un 6,25% a un 15,28%, aumento estadísticamente significativo.

3. En todos los grupos de edad y en ambos sexos el aumento ha sido constante en este periodo, y en el 2017, último año de la observación, se evidencian las cifras de prevalencia más altas en todos los nichos por edad y sexo considerados en comparación con los años precedentes.

4. Otra de las percepciones comunes, es que la alergia es una enfermedad de gente joven pero aquí presentamos datos cuantitativos de forma irrefutable. El grupo de 15-24 años destaca de forma uniforme en todos los años siendo el nicho al que afecta con más frecuencia la alergia. Como es lógico en los grupos de edad más avanzados se observa un claro descenso de afectación por alergia, con un dibujo en escalera descendente.

5. La alergia afecta en mayor medida a las mujeres y nuestros datos numéricos muestran claramente la diferencia a favor de este colectivo en cuanto a prevalencia y esto se da en todos los años del periodo.

6. Si tenemos en cuenta los niveles de formación alcanzados por la población, podemos ver que los valores de alergia son más altos en personas con niveles superiores formativos, tanto en cifras totales como si desglosamos en hombres y

mujeres, ocurriendo este hecho independientemente del año en que se analice. En resumen, destacamos que las cifras más bajas de prevalencia coinciden con el grupo de personas sin estudios y que las más altas son para grupos con niveles superiores cursados.

7. Respecto a la clasificación por niveles de renta, nuestros resultados muestran valores superiores de alergia, con significación estadística, de unos dos puntos en los niveles de renta más altos con respecto a los niveles de renta más bajos.

8. Si estudiamos la prevalencia de asma separada del resto de enfermedades alérgicas hay que destacar que no ha sufrido la misma evolución. Así, las cifras de asma se han mantenido estables en los últimos 12 años. Los máximos están en las poblaciones de más edad tanto en hombres como mujeres. El grupo de 75 y más años tiene la prevalencia más alta, en edades medias decae y a los 15-24 años asciende pero con menor intensidad que en edades avanzadas. Esta distribución en "V" o forma de valle se repite tanto en hombres como en mujeres. También destaca en el asma un marcado aumento en la población sin estudios, tanto para ambos sexos como por separado, y también de forma contraria a la alergia los niveles económicos más altos muestran bajos valores de asma.

9. Con respecto a la idea generalizada de que la mayor contaminación y el calentamiento global producen un aumento de la alergia, en nuestro caso si que se aprecia un aumento en las poblaciones de más de 400.000 habitantes donde se presupone una mayor contaminación. Pero si comparamos con el aumento global de alergia, independientemente del tipo de población, que ha sido superior a un 50% podríamos pensar que podría deberse a otros factores, como que los grupos de edad más jóvenes, y que tienen una mayor proporción de alergia, son más abundantes en las grandes ciudades. No existen diferencias significativas en el total de ambos sexos en asmáticos según población.

10. La Comunidad Valenciana mostraría un moderado aumento tanto en lo que respecta a la alergia como al asma. La disminución de prevalencia se ha consolidado

en Extremadura y Canarias tanto en alergia como en asma. Asturias no varía en alergia y Castilla-La Mancha no varía en asma ni en alergia.

A destacar en sentido negativo, el aumento en Melilla tanto de alergia como de asma y también la consolidación de Galicia en el aumento de ambas patologías. La cornisa mediterránea comprendiendo Cataluña y Murcia (sin la Comunidad Valenciana) más Baleares disminuirían la prevalencia de asma en la serie temporal tomada en el estudio.

11. La aportación de nuestro trabajo con técnicas multivariantes para el estudio de la epidemiología de la alergia y asma, destaca con la obtención de resultados mediante el análisis de componentes principales y el análisis de conglomerados. El análisis multivariante en su versión estadística y su aplicación gráfica permite una visión diferente de múltiples datos clínicos para poder sacar consecuencias prácticas.

En el caso de los conglomerados en la alergia, se ha conseguido distinguir entre grupos por medio de su porcentaje de prevalencia en muchos de los casos. Sin embargo, en el caso del análisis de componentes principales, en su versión gráfica del círculo de correlaciones, los datos no son relevantes para obtener una rápida visualización del tipo de grupos y de las consecuencias que podemos extraer. Esto se explica porque las variables se correlacionan positivamente y existe poco ángulo entre ellas con una agrupación abigarrada.

En lo referente al análisis de componentes principales en el asma, en su versión gráfica en el círculo de correlaciones es evidente la diferencia de distribución que veíamos con respecto a la alergia. La longitud de los vectores no es similar por lo que las variables muestran diferencias entre ellas. Las variables están poco correlacionadas porque existe un mayor ángulo entre ellas comparadas con la representación de la alergia, consiguiéndose en este caso una cierta diferenciación por grado de prevalencia.

12. El análisis multivariante también nos han permitido establecer interesantes relaciones con otros tipos de factores de salud, como es el tipo de trabajo en relación con el esfuerzo físico, el sedentarismo, el hábito tabáquico, el normopeso/

sobrepeso. Por otro lado, se ha ampliado la visión de la estadística descriptiva para conseguir una mayor visualización y comprensión de la misma.

En los registros temporales escogidos, la alergia se ha relacionado sobre todo con sedentarismo tanto en el ocio como en el trabajo. Mientras que el asma requiere actividad física moderada o algún grado de esfuerzo físico en el trabajo.

VII. BIBLIOGRAFÍA

1. Platts-Mills TA. The allergy epidemics: 1870–2010. *J Allergy Clin Immunol.* 2015; 136 (1):3–13.
2. Hill DA, Grundmeier RW, Ram G, Spergel JM. The epidemiologic characteristics of healthcare provider-diagnosed eczema, asthma, allergic rhinitis, and food allergy in children: a retrospective cohort study. *BMC Pediatr.* 2016; 16:133.
3. Smith DH, Malone DC, Lawson KA, Ikamoto J, Battista C, Saunders WB. A national estimate of the economic costs of asthma. *Am J Respir Crit Care Med.* 1997; 156: 787-93.
4. Cohen S, Dworetzky M, Frick OL. Coca and Cooke on the classification of hypersensitiveness. *J Allergy Clin Immunol.* 2003; 111 (1):205-210.
5. Siracusa MC, Kim BS, Spergel JM, Artis D. Basophils and allergic inflammation. *J Allergy Clin Immunol.* 2013; 132 (4):789–801.
6. Marenholz I, Esparza-Gordillo J, Lee YA. The genetics of the skin barrier in eczema and other allergic disorders. *Curr Opin Allergy Clin Immunol.* 2015; 15 (5):426–34.
7. Savenije OE, Mahachie John JM, Granell R, Kerkhof M, Dijk N, De Jongste JC, Smit HA, Brunekreef B, Postma DS, Van Steen K, Henderson J, Koppelman GH. Association of IL33-IL-1 receptorlike 1 (IL1RL1) pathway polymorphisms with wheezing phenotypes and asthma in childhood. *J Allergy Clin Immunol.* 2014; 134 (1):170–7.
8. Harada M, Hirota T, Jodo AI, Hitomi Y, Sakashita M, Tsunoda T, Miyagawa T, Doi S, Kameda M, Fujita K, Miyatake A, Enomoto T, Noguchi E, Masuko H, Sakamoto T, Hizawa N, Suzuki Y, Yoshihara S, Adachi M, Ebisawa M, Saito H, Matsumoto K, Nakajima T, Mathias RA, Rafaels N, Barnes KC, Himes BE, Duan QL, Tantisira KG, Weiss ST, Nakamura Y, Ziegler SF, Tamari M. Thymic stromal lymphopoietin gene promoter polymorphisms are associated with susceptibility to bronchial asthma. *Am J Respir Cell Mol Biol.* 2011; 44 (6):787–93.
9. Cookson W. The Alliance of genes and environment in asthma and allergy. *Nature.* 1999; 402 (6760 Suppl): B5-11.

10. Till S, Durham S, Dickason R, Houston D, Bungre J, Walker S, Robinson D, Kay AB, Corrigan C. IL-13 production by allergen-stimulated T cells is increased in allergic disease and associated with IL-5 but not IFN-gamma expression. *Immunology*. 1997; 91 (1): 53-7.
11. Romagnani S. Human TH1 and TH2 subsets: doubt no more. *Immunol Today*. 1991;12 (8): 256-7.
12. Constant SL, Bottomly K. Induction of Th1 and Th2 CD4 T cell responses: the alternative approaches. *Annu Rev Immunol*. 1997; 15 (1): 297-322.
13. Prescott S, Macaubas C, Holt BJ, Smallacombe TB, Loh R, Sly PD, Holt PG. Transplacental priming of the human immune system to environmental allergens: universal skewing of initial T cell responses towards the Th2 cytokine profile. *J Immunol*. 1998; 160 (10): 4730-7.
14. Hill DA, Spergel JM. The atopic march Critical evidence and clinical relevance. *Ann Allergy Asthma Immunol*. 2018; 120 (2): 131-7.
15. Kashem SW, Haniffa M, Kaplan DH. Antigen-presenting cells in the skin. *Annu Rev Immunol*. 2017; 35: 469-99.
16. Worbs T, Hammerschmidt SI, Förster R. Dendritic cell migration in health and disease. *Nat Rev Immunol*. 2017; 17(1): 30-48.
17. Noval Rivas M, Chatila TA. Regulatory T cells in allergic diseases. *J Allergy Clin Immunol*. 2016; 138(3):639-52.
18. Tait Wojno ED, Artis D. Emerging concepts and future challenges in innate lymphoid cell biology. *J Exp Med*. 2016; 213 (11):2229-48.
19. Galli SJ, Tsai M, Piliponsky AM. The development of allergic inflammation. *Nature*. 2008;454(7203):445-54.
20. Schatz M, Rosenwasser L. The allergic asthma phenotype. *J Allergy Clin Immunol Pract*. 2014; 2(6): 645-48.

21. Robinson DS, Hamid Q, Ying S, Tsiocopoulos A, Barkans J, Bentley AM, Corrigan C, Durham SR, Kay AB. Predominant Th2-type bronchoalveolar lavage T-lymphocyte population in atopic asthma. *N Engl J Med.* 1992; 326 (5): 298-304.
22. Silver JD, Spriggs K, Haberle S, Katelaris CH, Newbiggin EJ, Lampugnani ER. Crowd sourced allergic rhinitis symptom data: the influence of environmental and demographic factors. *Sci Total Environ.* 2020; 705: 135147.
23. Von Mutius E, Martinez FD, Fritzsich C, Nicolai T, Reitmeir P, Thiemann HH. Differences in the prevalence of asthma and atopic sensitisation between East and West Germany. *Am J Respir Crit Care Med.* 1994; 149: 358-64.
24. Von Mutius E, Weiland S, Fritzsich C, Duhme H, Keil U. Increasing prevalence of hay fever and atopy among children in Leipzig, East Germany. *Lancet.* 1998; 351 (9106): 862-6.
25. Magnussen H, Jörres R, Nowak D. Effect of air pollution on the prevalence of asthma and allergy: lessons from the German reunification. *Thorax.* 1993; 48(9): 879-81.
26. Strachan DP. Hay fever, hygiene and house-hold size. *BMJ.* 1989; 299 (6710):1259-60.
27. Sepp E, Julge K, Vasar M, Naaber P, Bjorksten B, Mikkelsaar M. Intestinal microflora of Estonian and Swedish infants. *Acta Paediatr.* 1997; 86 (9): 956-61.
28. Bjorksten B, Naaber P, Sepp E, Mikkelsaar M. The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin Exp Allergy.* 1999; 29 (3): 342-6.
29. Braun-Fahrländer C, Gassner M, Grize L, Neu U, Sennhauser FH, Varonier HS, Vuille JC, Wüthrich B. Prevalence of hay fever and allergic sensitization in farmers children and their peers living in the same rural community. *Clin Exp Allergy.* 1999; 29 (1): 28-34.
30. Stein MM, Hrusch CL, Gozdz J, Igartua C, Pivniouk V, Murray SE, Ledford JG, Marques Dos Santos M, Anderson RL, Metwali N, Neilson JW, Maier RM, Gilbert JA, Holbreich M, Thorne PS, Martinez FD, von Mutius E, Vercelli D, Ober C, Sperling AI.

Innate immunity and asthma risk in Amish and Hutterite farm children. *N Engl J Med.* 2016; 375 (5): 411-21.

31. Openshaw PJM, Hewitt CR. Protective and harmful effects of viral infections in childhood on wheezing disorders and asthma. *Am J Respir Crit Care Med.* 2000; 162 (2 Pt 2): S40-3.

32. Folkerts G, Walz G, Openshaw PMJ. Do common childhood infections teach the immune system not to be allergic?. *Immunol today.* 2000; 21 (3):118-20.

33. Burr ML. Pollution –does it cause asthma?. *Arch Dis Child.* 1995; 72 (5): 377-9.

34. Shultz AA, Schauer JJ, Malecki KM. Allergic disease associations with regional and localized estimates of air pollution. *Environ Res.* 2017; 155: 77-85.

35. Seaton A, Godden DJ, Brown K. Increase in asthma: a more toxic environment or a more susceptible population?. *Thorax.* 1994; 49 (2): 171-4.

36. Peat JK. The rising trend in allergic illness: which environmental factors are important?. *Clin Exp Allergy.* 1994; 24 (9): 797-800.

37. Samet JM, Marbury MC, Spengler JD. Health effects and sources of indoor air pollution. *Am Rev Respir Dis.* 1987;136 (6): 1486-508 and 1988; 137 (1): 221-42.

38. Baxi SN, Sheehan WJ, Sordillo JE, Muilenberg ML, Rogers CA, Gaffin JM, Permaul P, Lai PS, Louisias M, Petty CR, Fu C, Gold DR, Phipatanakul W. Association between fungal spore exposure in inner-city schools and asthma morbidity. *Ann Allergy Asthma Immunol.* 2019; 122 (6): 610-15.

39. Tham R, Erbas B, Dharmage SC, Tang ML, Aldakheel F, Lodge CJ, Thomas PS, Taylor PE, Abramson MJ, Lowe AJ. Outdoor fungal spores and acute respiratory effects in vulnerable individuals. *Environ Res.* 2019; 178: 108675.

40. Denning DW, Pashley C, Hartl D, Wardlaw A, Godet C, Del Giacco S, Delhaes L, Sergejeva S. Fungal allergy in asthma– state of the art and research needs. *Clin Transl Allergy.* 2014; 4: 14.

41. Fisk WJ, Lei-Gomez Q, Mendell MJ. Meta-analyses of the associations of respiratory health effects with dampness and mold in homes. *Indoor Air*. 2007; 17 (4): 284-96.
42. Stewart GA. Dust mite allergens. *Clin Rev Allergy Immunol*. 1995; 13 (2): 135-50.
43. Platts-Mills TA, Vaughan JW, Carter MC, Woodfolk JA. The role of intervention in established allergy: avoidance of indoor allergens in the treatment of chronic allergic disease. *J Allergy Clin Immunol*. 2000; 106 (5):787-804.
44. Bates DV, Baker-Anderson M, Sizto R. Asthma attack periodicity: a study of hospital emergency visits in Vancouver. *Environ Res*. 1990; 51 (1): 51-70.
45. Thurston GD, Ito K, Kinney PL, Lippmann M. A multi-year study of air pollution and respiratory hospital admissions in three New York State metropolitan areas: results for 1988 and 1989 summers. *J Exp Anal Environ Epidemiol*. 1992; 2 (4): 429-50.
46. Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis*. 1993; 147 (4): 826-31.
47. Anderson HR, Ponce de Leon A, Bland JM, Bower JS, Emberlin J, Strachan DP. Air pollution, pollution, pollens and daily admissions for asthma in London 1987-92. *Thorax*. 1998; 53 (10): 842-48.
48. Cakmak S, Dales RE, Burnett RT, Judek S, Coates F, Brook JR. Effect of airborne allergens on emergency visits by children for conjunctivitis and rhinitis. *Lancet*. 2002; 359 (9310): 947-8.
49. Sedghy F, Varasteh AR, Sankian M, Moghadam M. Interaction Between Air Pollutants and Pollen Grains: The Role on the Rising Trend in Allergy. *Rep Biochem Mol Biol*. 2018; 6 (2):219-24.
50. Nel AE, Diaz-Sanchez D, Li N. The role of particulate pollutants in pulmonary inflammation and asthma: evidence for the involvement of organic chemicals and oxidative stress. *Curr Opin Pulm Med*. 2001; 7 (1):20-6.

51. Chung HY, Hsieh CJ, Tseng CC, Yiin LM. Association between the first occurrence of allergic rhinitis in preschool children and air pollution in Taiwan. *Int J Environ Res Public Health*. 2016; 13 (3):268.
52. Takano H, Inoue KI. Environmental pollution and allergies. *J Toxicol Pathol*. 2017; 30 (3): 193-9.
53. Eguiluz-Gracia I, Mathioudakis AG, Bartel S, Vijverberg SJH, Fuertes E, Comberati P, Cai YS, Tomazic PV, Diamant Z, Vestbo J, Galan C, Hoffmann B. The need for clean air: the way air pollution and climate change affect allergic rhinitis and asthma. *Allergy*. 2020; 75 (9): 2170-2184.
54. Abbey DE, Petersen F, Mills PK, Beeson WL. Long-term ambient concentrations of total suspended particulates, ozone and sulfur dioxide and respiratory symptoms in a non-smoking population. *Arch Environ Health*. 1993; 48 (1): 33-47.
55. Paterson CA, Sharpe RA, Taylor T, Morrissey K. Indoor PM_{2.5}, VOCs and asthma outcomes: a systematic review in adults and their home environments. *Environ Res*. 2021; 202:111631.
56. Baldacci S, Maio S, Cerrai S, Sarno G, Baiz N, Simoni M, Annesi-Maesano I, Viegi G; HEALS Study. Allergy and asthma: Effects of the exposure to particulate matter and biological allergens. *Respir Med*. 2015;109 (9):1089-104.
57. Delfino RJ, Becklake MR, Hanley JA. The relationship of urgent hospital admissions for respiratory illnesses to photochemical air pollution levels in Montreal. *Environ Res*. 1994;67 (1): 1-19.
58. Cody RP, Weisel CP, Birnbaum G, Liroy PJ. The effect of ozone associated with summertime photochemical smog on the frequency of asthma visits to hospital emergency departments. *Environ Res*. 1992; 58 (2): 184-94.
59. Filippidou EC, Koukouliata A. Ozone effects on the respiratory system. *Prog Health Sci*. 2011; 1:144-55.
60. D Amato M, Cecchi L, Annesi-Maesano I, D Amato G. News on Climate Change, Air Pollution, and Allergic Triggers of Asthma. *J Investig Allergol Clin Immunol*. 2018;28 (2):91-7.

61. Von Mutius E, Martinez FD, Fritzsche C, Nicolai T, Roeli G, Thiemann HH. Prevalence of asthma and atopy in two areas of West and East Germany. *Am J Respir Crit Care Med.* 1994; 149 (2 Pt 1):358-64.
62. Nowak D, Heinrich J, Jörres R, Wassmer G, Berger J, Beck E, Boczor S, Claussen M, Wichmann HE, Magnussen H. Prevalence of respiratory symptoms, bronchial hyperresponsiveness and atopy among adults: West and East Germany. *Eur Respir J.* 1996; 9 (12): 2541-52.
63. Wjst M, Reitmeir P, Dold S, Wulff A, Nicolai T, von Loeffelholz-Colberg EF, von Mutius E. Road traffic and adverse effects on respiratory health in children. *BMJ.* 1993; 307 (6904):596-600.
64. Livingstone AE, Shaddick G, Grundy C, Elliott P. Do people living near inner city main roads have more asthma requiring treatment? Case control study. *BMJ.* 1996; 312 (7032): 676-7.
65. Oosterlee A, Drijver M, Lebret E, Brunekreef B. Chronic respiratory symptoms in children and adults living along streets with high traffic density. *Occup Environ Med.* 1996; 53 (4): 241-7.
66. Van Vliet P, Knape M, De Hartog J, Janssen N, Harssema H, Brunekreef B. Motor vehicle exhaust and chronic respiratory symptoms in children living near freeways. *Environ Res.* 1997; 74 (2): 122-32.
67. Weiland SK, Mundt KA, Rückmann A, Keil U. Self-reported wheezing and allergic rhinitis in children and traffic density on street of residence. *Ann Epidemiol.* 1994; 4 (3): 243-7.
68. Wilkinson P, Elliott P, Grundy C, Shaddick G, Thakrar B, Walls P, Falconer S. Case-control study of hospital admission with asthma in children aged 5-14 years: relation with road traffic in north west London. *Thorax.* 1999; 54 (12): 1070-4.
69. Gauderman WJ, Vora H, McConnell R, Berhane K, Gilliland F, Thomas D, Lurmann F, Avol E, Kunzli N, Jerrett M, Peters J. Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study. *Lancet.* 2007; 369 (9561):571-7.

70. Yanagisawa R, Takano H, Inoue KI, Ichinose T, Sadakane K, Yoshino S, Yamaki K, Yoshikawa T, Hayakawa K. Components of diesel exhaust particles differentially affect Th1/Th2 response in a murine model of allergic airway inflammation. *Clin Exp Allergy*. 2006; 36 (3): 386-95.
71. Carlsten C, Blomberg A, Pui M, Sandstrom T, Wong SW, Alexis N, et al. Diesel exhaust augments allergen induced lower airway inflammation in allergic individuals: a controlled human exposure study. *Thorax*. 2016; 71 (1):35-44.
72. Marks GB, Bush RK. It's blowing in the wind: new insights into thunderstorm-related asthma. *J Allergy Clin Immunol*. 2007; 120 (3): 530-32.
73. Price D, Hughes KM, Thien F, Suphioglu C. Epidemic thunderstorm asthma: lessons learned from the storm down-under. *J Allergy Clin Immunol Pract*. 2021; 9 (4): 1510-15.
74. Harun NS, Lachapelle P, Douglass J. Thunderstorm-triggered asthma: what we know so far. *J Asthma Allergy*. 2019; 12: 101-8.
75. Bellomo R, Gigliotti P, Treloar A, Holmes P, Suphioglu C, Singh MB, Knox B. Two consecutive thunderstorm associated epidemics of asthma in the city of Melbourne. The possible role of rye grass pollen. *Med J Aust*. 1992; 156 (12): 834-7.
76. Venables KM, Allitt U, Collier CG, Emberlin J, Greig JB, Hardaker PJ, Highham JH, Laing-Morton T, Maynard RL, Murray V, Strachan D, Tee RD. Thunderstorm-related asthma: the epidemic of 24/25 June 1994. *Clin Exp Allergy*. 1997; 27 (7): 725-36.
77. Knox RB. Grass pollen, thunderstorms and asthma. *Clin Exp Allergy*. 1993; 23 (5): 354-9.
78. Newson R, Strachan D, Archibald E, Emberlin J, Hardaker P, Collier C. The effect of thunderstorms and airborne grass pollen on acute asthma incidence in England, 1990-1994. *Thorax*. 1997; 52 (8): 680-5.
79. Bowatte G, Lodge CJ, Knibbs LD, Lowe AJ, Erbas B, Dennekamp M, Marks GB, Giles G, Morrison S, Thompson B, Thomas PS, Hui J, Perret JL, Abramson MJ, Walters H, Matheson MC, Dharmage SC. Traffic-related air pollution exposure is associated

with allergic sensitization, asthma, and poor lung function in middle age. *J Allergy Clin Immunol.* 2017; 139 (1):122-9.

80. Beggs P. Climate change and allergy in Australia: an innovative, high-income country, at potential risk. *Public Health Res Pract.* 2018; 28 (4): 2841828.

81. Timm S, Frydenberg M, Janson C, Campbell B, Forsberg B, Gislason T, Holm M, Jogi R, Omenaas E, Sigsgaard T, Svanes C, Schlünssen V. The urban-rural gradient in asthma: a population-based study in northern Europe. *Int J Environ Res Public Health.* 2015; 13 (1):93.

82. N Newson R, Strachan D, Archibald E, Emberlin J, Hardaker P, Collier C. Acute asthma epidemics and their predictors in England, 1987-1994. *Eur Respir J.* 1998; 11: 694-701.

83. Burr ML, Butland BK, King S, Vaughan Williams E. Changes in asthma prevalence: two surveys 15 years apart. *Arch Dis Child.* 1989; 64 (10): 1452-6.

84. Johnston SL, Pattemore PK, Sanderson G, Smith S, Lampe F, Josephs L, Symington P, O'Toole S, Myint SH, Tyrrell DA. Community study of the role of viral infections in exacerbations of asthma in 9-11-year-old children. *BMJ.* 1995; 310 (6989): 1225-9.

85. Nurmagambetov T, Kuwahara R, Garbe P. The economic burden of asthma in the United States, 2008-2013. *Ann Am Thorac Soc.* 2018;15 (3):348-56.

86. Asher MI, Pattemore PK, Harrison AC, Mitchell EA, Rea HH, Stewart AW, Woolcock AJ. International comparison of the prevalence of asthma symptoms and bronchial hyperresponsiveness. *Am Rev Respir Dis.* 1988; 138 (3): 524-9.

87. Johannessen A, Verlato G, Benediktsdottir B, Forsberg B, Franklin K, Gislason T, Holm M, Janson C, Jögi R, Lindberg E, Macsali F, Omenaas E, Real FG, Saure EW, Schlünssen V, Sigsgaard T, Skorge TD, Svanes C, Torén K, Waatevik M, Nilsen RM, de Marco R. Longterm follow-up in European respiratory health studies - patterns and implications. *BMC Pulm Med.* 2014; 14:63.

88. Morgan WJ, Stern DA, Sherrill DL, Guerra S, Holberg CJ, Guilbert TW, Taussig LM, Wright AL, Martinez FD. Outcome of asthma and wheezing in the first 6 years of life: follow-up through adolescence. *Am J Respir Crit Care Med.* 2005; 172 (10):1253-8.

89. Thorn J, Brisman J, Torén K. Adult-onset asthma is associated with self-reported mold or environmental tobacco smoke exposures in the home. *Allergy*. 2001;56 (4): 287-92.
90. Vargas Becerra MH. Epidemiología del asma. *Rev Alert Mex*. 2009;56(Suppl 1): S3-9.
91. Settipane RA, Charnock DR. Epidemiology of rhinitis: allergic and nonallergic. *Clin Allergy Immunol*. 2007; 19:23-34.
92. Eifan AO, Durham SR. Pathogenesis of rhinitis. *Clin Exp Allergy*. 2016;46(9):1139-51.
93. Williams HC. Atopic dermatitis. *N Engl J Med*. 2005; 352 (22):2314-24.
94. Novak N, Bieber T. Allergic and nonallergic forms of atopic diseases. *J Allergy Clin Immunol*. 2003; 112 (2): 252-62.
95. Sicherer SH, Sampson HA. Food allergy. *J Allergy Clin Immunol*. 2006; 117(2 Suppl Mini-Primer) S:470-5.
96. Branum AM, Lukacs SL. Food allergy among children in the United States. *Pediatrics*. 2009; 124 (6):1549-55.
97. Esbo M, Botten G, Hiborsen R, Magnus P. The prevalence of allergy to egg: a population based study in young children. *Allergy*. 2001; 56 (5):403-11.
98. Bock SA. Prospective appraisal of complaints of adverse reactions to foods in children during the first 3 years of life. *Pediatrics*. 1987; 79 (5):683-8.
99. Mullins RJ. Anaphylaxis: risk factors for recurrence. *Clin Exp Allergy*. 2003; 33 (8):1033-40.
100. Zuberbier T, Balke M, Worm M, Edenharter G, Maurer M. Epidemiology of urticaria: a representative cross-sectional population survey. *Clin Exp Dermatol*. 2010; 35 (8):869-73.
101. Kaplan AP. Chronic urticarial: pathogenesis and treatment. *J Allergy Clin Immunol*. 2004; 114 (3):465-74.

102. Confino-Cohen R, Chodick G, Shalev V, Leshno M, Kimhi O, Golodberg A. Chronic urticarial and autoimmunity: associations found in a large population study. *J Allergy Clin Immunol*. 2012; 129(5):1307-13.
103. Grattan CEH, Humphreys F. Guidelines for evaluation and management of urticarial in adults and children. *Br J Dermatol*. 2007; 157 (6):1116-23.
104. Simoens S, Laekkeman G. Pharmacotherapy of allergic rhinitis: a pharmacoeconomic approach. *Allergy*. 2009; 64 (1):85-95.
105. Martínez-Moragón E, Serra-Batlles J, De Diego A. Economic cost of treating the patient with asthma in Spain: The asthma cost study. *Arch Bronconeumol*. 2009; 45(10): 481-6.
106. Cockburn IM, Bailit HL, Berndt ER, Finkelstein SN. Loss of work productivity due to illness and medical treatment. *J Occup Environ Med*. 1999; 41 (11): 948-53.
107. Ojeda P, Sanz de Burgoa V. Costs associated with workdays lost and utilization of health care resources because of asthma in daily clinical practice in Spain. *J Investig Allergol Clin Immunol*. 2013;23 (4): 234-41.
108. Bousquet JA. Physicians view of health economics in asthma. *Eur Resp Rev*. 1995; 5:275-8.
109. Szucs TD, Anderhub H, Rutishauser M. The economic burden of asthma: direct and indirect costs in Switzerland. *Eur Respir J*. 1999; 13(2):281-6.
110. Alarcón J. Epidemiología: concepto, usos y perspectivas. *Rev Perú epidemiol*. 2009;13(1).
111. Rodríguez Artalejo F, Ortún Rubio V, Banegas JR, Martín Moreno JM. La epidemiología como instrumento para una política de salud racional. *Med Clin (Barc)*. 1989; 93 (17): 663-6.
112. Hernández-Ávila M, Garrido-Latorre F, López-Moreno S. Diseño de estudios epidemiológicos. *Salud Pública Mex*. 2000; 42(2):144-54.
113. Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am J Epidemiol*. 1991; 133 (7): 635-48.

114. Restrepo Chavarriaga G, Gonzalez Quiñones JC. Libro de biometría comunitaria. Bogotá, Colombia: Fundación Universitaria Juan N Corpas; 2010.
115. Argimon JP, Jiménez JV. Métodos de investigación clínica y epidemiológica. 3a ed. Madrid, España: Elsevier España; 2004.
116. Greemberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. Medical epidemiology. 3a ed. New York, USA: McGraw-Hill; 2001.
117. Londoño JLF. Metodología de la investigación epidemiológica. 3a ed. Bogotá, Colombia: El Manuel Moderno; 2004.
118. Ahlbom A, Staffan N. Fundamentos de epidemiología. 4a ed. México: Siglo XXI; 1993.
119. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002; 359 (9302):248-52.
120. Ulin PR, Robinson ET, Tolley EE. Investigación aplicada en salud pública: métodos cualitativos. Washington DC, USA: Organización Panamericana de la Salud; 2006.
121. Parascandola M, Weed DL. Causation in epidemiology. *J Epidemiol Community Health*. 2001; 55 (12):905-12.
122. Grimes DA, Schulz KF. Descriptive studies: What they can and cannot do. *Lancet*. 2002; 359 (9301):145-9.
123. Manterola C, Otzen T. Estudios observacionales. los diseños utilizados con mayor frecuencia en investigación clínica. *Int J Morphol*. 2014; 32(2):634-45.
124. Walter SD. The ecologic method in the study of environmental health. II. Methodologic issues and feasibility. *Environ Health Perspect*. 1991; 94:67-73.
125. Manterola C. Estrategias de investigación. Diseños observacionales 1a parte. Estudios descriptivos. *Rev Chil Cir*. 2001; 53:229-33.
126. Groenwold RHH, Hoes AW. Confounding in publications of observational intervention studies. *Eur J Epidemiol*. 2007; 22 (7): 413-5.

127. Manterola C. Estrategias de investigación. Diseños observacionales 3a parte. Estudio de cohortes. *Rev Chil Cir.* 2001; 53:410-4.
128. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ.* 1996; 312 (7040):1215-8.
129. García-de la Torre GS, Huerta-Alvarado SG. Consideraciones metodológicas y análisis simple de los estudios transversales. *Bol Med Hosp Infant Mex.* 1998; 55 (6):348-56.
130. Schiaffino A, Rodríguez M, Pasarín MI, Regidor E, Borrell C, Fernández E. ¿Odds ratio o razón de proporciones?. Su utilización en estudios transversales. *Gac Sanit.* 2003; 17(1): 70-4.
131. Lee J. Odds ratio or relative risk for cross-sectional data?. *Int J Epidemiol.* 1994; 23 (1): 201-3.
132. Levin KA. Study design III: Cross-sectional studies. *Evid Based Dent.* 2006; 7 (1): 24-5.
133. Lasky T, Stolley PD. Selection of cases and controls. *Epidemiol Rev.* 1994; 16 (1): 6-17.
134. Moreno-Altamirano L, Cano-Valle F, García-Romero H. Epidemiología clínica. México: McGraw-Hill Interamericana; 1994.
135. Lascano-Ponce E, Salazar-Martínez E, Hernández-Ávila M. Estudios epidemiológicos de casos y controles. Fundamento teórico, variantes y aplicaciones. *Salud Publica Mex.* 2001; 43 (2):135-50.
136. Morris JN. Uses of epidemiology. *Br Med J.* 1955; 2 (4936):395-401.
137. Robinson WS. Ecological correlations and the behavior of individuals. *Int J Epidemiol.* 2009; 38 (2): 337-41.
138. Lynch SM, Rusiecki JA, Blair A, Dosemeci M, Lubin J, Sandler D, Hoppin JA, Lynch CF, Alavanja MC. Cancer incidence among pesticide applicators exposed to cyanazine in the agricultural health study. *Environ Health Perspect.* 2006; 114 (8):1248-52.

139. Conceição GM, Miraglia SG, Kishi HS, Saldiva PH, Singer JM. Air pollution and child mortality: A time-series study in São Paulo, Brazil. *Environ Health Perspect.* 2001;109(Suppl 3):347-50.
140. Forget G, Lebel J. An ecosystem approach to human health. *Int J Occup Environ Health.* 2001;7(2 Suppl): S3-38.
141. Trickett EJ, Beehler S, Deutsch C, Green LW, Hawe P, McLeroy K, Miller RL, Rapkin BD, Schensul JJ, Schulz AJ, Trimble JE. Advancing the science of community level interventions. *Am J Public Health.* 2011; 101 (8):1410-9.
142. Morgenstern H. Ecologic studies in epidemiology: Concepts, principles, and methods. *Annu Rev Public Health.* 1995; 16:61-81.
143. Wakefield J. Ecologic studies revisited. *Annu Rev Public Health.* 2008; 29:75-90.
144. Cronon W. The uses of environmental history. *Environ Hist Rev.* 1993; 17 (3): 1-22.
145. Dufault B, Klar N. The quality of modern cross-sectional ecologic studies: A bibliometric review. *Am J Epidemiol.* 2011; 174 (10):1101-7.
146. Brown P. Qualitative methods in environmental health research. *Environ Health Perspect.* 2003; 111 (14):1789-98.
147. Scammell MK. Qualitative environmental health research: An analysis of the literature, 1991-2008. *Environ Health Perspect.* 2010; 118 (8):1146-54.
148. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health.* 1982; 72 (12):1336-44.
149. Nicholson R. Environmental health practice: For today and for the future. *Environ Health.* 2001; 1 (3):73-84.
150. Mora AD, Chamizo GH, Mata SA. Estudio exploratorio sobre la incidencia de cáncer gástrico y los contenidos de nitratos en el agua potable en Costa Rica. *Rev Costarric Salud Pública.* 2006; 15 (28):17-28.

151. Brown P. Popular epidemiology and toxic waste contamination: Lay and professional ways of knowing. *J Health Soc Behav.* 1992; 33 (3):267-81.
152. Idrovo AJ. Plaguicidas usados en la fumigación de cultivos ilícitos y salud humana: ¿una cuestión de ciencia o política?. *Rev Salud Pública (Bogotá).* 2004; 6 (2):199-211.
153. Hernández Ávila M, Garrido Latorre F, López Moreno S. Diseño de estudios epidemiológicos. *Salud Publica Mex.* 2000; 42 (2):144-54.
154. Méndez Paz F. El papel de la Epidemiología Ambiental en el Desarrollo Disciplinar de la Epidemiología. *Colomb Med.* 2011; 42 (3): 278-85.
155. Sackett D, Haynes B, Tugwell P. Epidemiología clínica: una ciencia básica para la medicina clínica. Madrid: Díaz de Santos; 1989.
156. Alergológica 1992: factores epidemiológicos, clínicos y socioeconómicos de las enfermedades alérgicas en España 1992. SEAIC; 1992.
157. Alergológica 2005: factores epidemiológicos, clínicos y socioeconómicos de las enfermedades alérgicas en España 2005. SEAIC; 2006.
158. Alergológica 2015: factores epidemiológicos, clínicos y socioeconómicos de las enfermedades alérgicas en España 2015. SEAIC; 2015.
159. Gaig P, García Abujeta JL, Muñoz-Lejarazu D, Leonart Bellfill R, Caballero T. Prevalencia de alergia en la población adulta Española. *Alergol Inmunol Clin.* 2004; 19(2):68-74.
160. Asher MI, Weiland SK. ISAAC Steering Committee. The International Study of Asthma and Allergies in Childhood (ISAAC). *Clin Exp Allergy.* 1998; 28 (Suppl 5): 52-66.
161. Strachan D, Sibbald B, Weiland S, Ait-Khaled N, Anabwani G, Anderson HR, Asher MI, Beasley R, Björkstén B, Burr M, Clayton T, Crane J, Ellwood P, Keil U, Lai C, Mallol J, Martinez F, Mitchell E, Montefort S, Pearce N, Robertson C, Shah J, Stewart A, von Mutius E, Williams H. Worldwide Variations in prevalence of symptoms of

allergic rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC). *Pediatr Allergy Immunol.* 1997; 8(4): 161-76.

162. Williams H, Robertson C, Stewart A, Ait-Khaled N, Anabwani G, Anderson HR, Asher MI, Beasley R, Björkstén B, Burr M, Clayton T, Crane J, Ellwood P, Keil U, Lai C, Mallol J, Martinez F, Mitchell E, Montefort S, Pearce N, Shah J, Sibbald B, Strachan D, von Mutius E, Weiland S. Worldwide variations in the prevalence of symptoms of atopic eczema in the international study of asthma and allergies in childhood. *J Allergy Clin Immunol.* 1999; 103(1 Pt 1): 125-38.

163. Weinmayr G, Weiland SK, Björkstén B, Brunekreef B, Büchele G, Cookson WO, García-Marcos L, Gotua M, Gratziau C, van Hage M, von Mutius E, Riiikjäv MA, Rzehak P, Stein RT, Strachan DP, Tsanakas J, Wickens K, Wong GW; ISAAC Phase Two Study Group. Atopic sensitization and the international variation of asthma symptom prevalence in children. *Am J Respir Crit Care Med.* 2007; 176(6): 565-74.

164. Ait-Khaled N, Pearce N, Anderson HR, Ellwood P, Montefort S, Shah J; ISAAC Phase Three Study Group. Global map of the prevalence of symptoms of rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three. *Allergy.* 2009; 64: 123-48.

165. Pearce N, Ait-Khaled N, Beasley R, Mallol J, Keil U, Mitchell E, Robertson C; ISAAC Phase Three Study Group. Worldwide trends in the prevalence of asthma symptoms: Phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax.* 2007; 62(9): 758-66.

166. Genuneit J, Cantelmo JL, Weinmayr G, Wong GWK, Cooper PJ, Riiikjäv MA, Gotua M, Kabesch M, von Mutius E, Forastiere F, Crane J, Nystad W, El Sharif N, Batllés-Garrido J, García-Marcos L, García-Hernández G, Morales Suárez-Varela MM, Nilsson L, Bråbäck L, Saraçlar Y, Weiland SK, Cookson WOC, Strachan DP, Moffatt MF; ISAAC Phase Two Study Group. A multi-centre study of candidate genes for wheeze and allergy. The International Study of Asthma and Allergies in Childhood Phase Two. *Clin Exp Allergy.* 2009; 39(12): 1875-88.

167. Flohr C, Weiland SK, Weinmayr G, Björkstén B, Bråbäck L, Brunekreef B, Büchele G, Clausen M, Cookson WO, von Mutius E, Strachan DP, Williams HC; ISAAC

Phase Two Study Group. The role of atopic sensitization in flexural eczema: findings from the International Study of Asthma and Allergies in Childhood Phase Two. *J Allergy Clin Immunol.* 2008;121(1):141-7.

168. Weinmayr G, Forastiere F, Weiland SK, Rzehak P, Abramidze T, Annesi-Maesano I, Björkstén B, Brunekreef B, Büchele G, Cookson WO, von Mutius E, Pistelli R, Strachan DP; ISAAC Phase Two Study Group. International variation in prevalence of rhinitis and its relation with sensitization to perennial and seasonal allergens. *Eur Respir J.* 2008; 32: 1250–1261.

169. Flohr C, Weinmayr G, Weiland SK, Addo-Yobo E, Annesi-Maesano I, Björkstén B, Bråbäck L, Büchele G, Chico M, Cooper P, Clausen M, El-Sharif N, Martínez Gimeno A, Mathur RS, von Mutius E, Morales Suárez-Varela MM, Pearce N, Svabe V, Wong GWK, Yu M, Zhong NS, Williams HC; ISAAC Phase Two Study Group. How well do questionnaires perform compared with physical examination in detecting flexural eczema? Findings from the International Study of Asthma and Allergies in Childhood (ISAAC) Phase Two. *Br J Dermatol.* 2009; 161(4): 846-853.

170. Stewart AW, Mitchell EA, Pearce N, Strachan DP, Weiland SK; ISAAC Steering Committee. The relationship of per capita gross national product to the prevalence of symptoms of asthma and other atopic diseases in children (ISAAC). *Int J Epidemiol.* 2001; 30: 173-9.

171. Anderson HR, Ruggles R, Pandey KD, Kapetanakis V, Brunekreef B, Lai CKW, Strachan DP, Weiland SK; ISAAC Phase One Study Group. Ambient particulate pollution and the world-wide prevalence of asthma, rhinoconjunctivitis and eczema in children: Phase One of the International Study of Asthma and Allergies in Childhood (ISAAC). *Occup Environ Med.* 2010; 67(5): 293-300.

172. Foliaki S, Kildegaard Nielsen S, Björkstén B, von Mutius E, Cheng S, Pearce N; ISAAC Phase One Study Group. Antibiotic sales and the prevalence of symptoms of asthma, rhinitis, and eczema: The International Study of Asthma and Allergies in Childhood (ISAAC). *Int J Epidemiol.* 2004; 33(3): 558-63.

173. Weiland SK, Hüsing A, Strachan D, Rzehak P, Pearce N; ISAAC Phase One Study Group. Climate and the prevalence of symptoms of asthma, allergic rhinitis and atopic eczema in children. *Occup Environ Med.* 2004; 61(7): 609-15.
174. Ellwood P, Asher MI, Björkstén B, Burr M, Pearce N, Robertson CF; ISAAC Phase One Study Group. Diet and asthma, allergic rhinoconjunctivitis and atopic eczema symptom prevalence: An ecological analysis of the International Study of Asthma and Allergies in Childhood (ISAAC) data. *Eur Respir J.* 2001; 17(3): 436-43.
175. Burr ML, Emberlin JC, Treu R, Cheng S, Pearce NE; ISAAC Phase One Study Group. Pollen counts in relation to the prevalence of allergic rhinoconjunctivitis, asthma and atopic eczema in the International Study of Asthma and Allergies in Childhood (ISAAC). *Clin Exp Allergy.* 2003; 33(12): 1675-80.
176. Mitchell EA, Stewart AW; ISAAC Phase One Study Group. The ecological relationship of tobacco smoking to the prevalence of symptoms of asthma and other atopic diseases in children: The International Study of Asthma and Allergies in Childhood (ISAAC). *Eur J Epidemiol.* 2001; 17(7): 667-73.
177. Von Mutius E, Pearce N, Beasley R, Cheng S, von Ehrenstein O, Björkstén B, Weiland S; ISAAC Steering Committee. International patterns of tuberculosis and the prevalence of symptoms of asthma, rhinitis and eczema. *Thorax.* 2000; 55(6): 449-53.
178. Shirtcliffe P, Weatherall M, Beasley R; ISAAC Phase One Study Group. An inverse correlation between estimated tuberculosis notification rates and asthma symptoms. *Respirology.* 2002; 7(2):153-5.
179. Pearce N, Ait-Khaled N, Beasley R, Mallol J, Keil U, Mitchell E, Robertson C; ISAAC Phase Three Study Group. Worldwide trends in the prevalence of asthma symptoms: Phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax.* 2007; 62(9):758-66.
180. Björkstén B, Clayton T, Ellwood P, Stewart A, Strachan D; ISAAC Phase Three Study Group. Worldwide time trends for symptoms of rhinitis and conjunctivitis:

Phase III of the International Study of Asthma and Allergies in Childhood. *Pediatr Allergy Immunol.* 2008; 19(2): 110-24.

181. Williams H, Stewart A, von Mutius E, Cookson B, Anderson HR; ISAAC Phase One and Three Study groups. Is eczema really on the increase worldwide?. *J Allergy Clin Immunol.* 2008; 121(4): 947-54.

182. Grupo ISAAC español. Prevalencia de síntomas sugestivos de rinitis alérgica y de dermatitis atópica en adolescentes (Estudio ISAAC España). *An Esp Pediatr.* 1999; 51:369-76.

183. García-Marcos Alvarez L, Martínez Torres A, Batlles Garrido J, Morales Suárez-Varela M, García Hernández G, Escribano Montaner A; Grupo ISAAC Español fase II. International Study of Asthma and Allergies in Childhood (ISAAC) fase II: metodología y resultados de participación en España. *An Esp Pediatr.* 2001; 55: 400-405.

184. Bercedo Sanz A, Martínez-Torres A, González Díaz C, López-Silvarrey Varela A, Pellegrini Belinchón FJ, Aguinaga-Ontoso I, García-Marcos L, Grupo GAN España. Prevalencia y evolución temporal de síntomas de asma en España. Estudio Global Asthma Network (GAN). *An Pediatr.* 2022; 97:161-71.

185. Chen M, Mao S, Liu Y. Big Data: A Survey. *Mob Networks Appl.* 2014; 19: 171–209.

186. Md Mujeeb S, Kasi Naidu L. A relative study on Big Data applications and techniques. *Int J Eng Innov Technol.* 2015;4 (10): 133-8.

187. Chen J, Chen Y, Du X, Li C, Lu J, Zhao S, Zhou X. Big data challenge: a data management perspective. *Front Comput Sci.* 2013; 7: 157-64.

188. Jin X, Wah BW, Cheng X, Wang Y. Significance and challenges of Big Data research. *Big Data Res.* 2015; 2: 59–64.

189. Jaramillo Valbuena S, Londoño JM. Sistemas para almacenar grandes volúmenes de datos. *Gerenc Tecnol Inform.* 2015; 13:17–28.

190. Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. *Int J Inf Manage.* 2015;35 (2):137-44.

191. Chen CP, Zhang CY. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Inform Sci.* 2014; 275: 314-47.
192. Barranco R. ¿Qué es Big Data? Todos formamos parte de ese gran crecimiento de datos. *Developer Works.* 2012;2-11.
193. Wu X, Zhu X, Wu GQ, Ding W. Data mining with big data. *IEEE Trans Knowl Data Eng.* 2014; 26: 97-107.
194. Domingos P. A few useful things to know about machine learning. *Commun ACM.* 2012; 55(10):78.
195. Portugal I, Alencar P, Cowan D. The use of machine learning algorithms in recommender systems: a systematic review. *Expert Syst Appl.* 2015; 4: 1-16.
196. Pérez Juste R. Estadística aplicada a las ciencias sociales. Madrid: UNED; 2012.
197. Meena MJ, Chandran KR, Karthik A, Samuel AV. An enhanced ACO algorithm to select features for text categorization and its parallelization. *Expert Syst Appl.* 2012; 39(5): 5861-587.
198. Tan M, Tsang IW, Wang L. Towards ultrahigh dimensional feature selection for big data. *J Mach Learn Res.* 2014;15(1):1371-429.
199. Fernández A, Del Río S, López V, Bawakid A, Del Jesús MJ, Benítez JM, Herrera F. Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2014; 4(5): 380-409.
200. Maté C. Big data. Un nuevo paradigma de análisis de datos. *Anales de mecánica y electricidad.* 2014; 91(6):10-6.
201. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst.* 2014; 2:3.
202. Fung K, Fang KT. Multiple outlier detection in multivariate data using projection pursuit techniques. *Journal of Statistical Planning and Inference.* 2000; 83(1):153-67.

203. Montanero Fernández J. Análisis multivariante. Universidad de Extremadura, Servicio de Publicaciones; 2008.
204. Uriel E, Aldas J. Análisis multivariante aplicado con R. Madrid: Ediciones Paraninfo; 2017.
205. Gutiérrez R, González A, Torres F, Gallardo JA. Técnicas de Análisis de Datos Multivariable. Tratamiento Computacional. Introducción a las Técnicas Factoriales y de Clasificación Automática. Universidad de Granada; 1994.
206. Cuadras CM. Nuevos métodos de análisis multivariante. Barcelona: CMC Editions; 2019.
207. Martínez Arias R. El análisis multivariante en la investigación científica. Madrid: Editorial La Muralla; 2008.
208. Ximénez MC, San Martín R. Fundamentos de las técnicas multivariantes. Madrid: UNED; 2004.
209. Escofier B, Pages J. Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación. Bilbao: UPV; 1992.
210. Arriaza AJ, Fernández F, Lopez MA, Muñoz M, Perez S, Sánchez A. Estadística Básica con R y R-Commander. Cádiz: Servicio de publicaciones de la Universidad de Cádiz; 2008.
211. Ministerio de Sanidad y Política Social. Encuesta Nacional de salud de España. <http://www.msps.es/estadEstudios/estadisticas/encuestaNacional/ense.htm>.
212. Ministerio de Sanidad y Política Social. Encuesta Nacional de salud de España. <http://www.msps.es/estadEstudios/encuestaNacional/añosanteriores.htm>.
213. Instituto Nacional de Estadística (INE). <http://ine.es/dyngs/INEbase/es>.
214. Carrasco JL. El método estadístico en la investigación médica. Madrid: Ciencia; 1996.
215. Cepeda Cuervo E. Estadística Matemática. Bogotá: Universidad Nacional de Colombia; 2008.

216. Ruiz Muñoz D. Manual de estadística. Eumed-Net; 2004.
217. Moore DS. Estadística aplicada básica. 2a ed. Antoni Bosch; 2004.
218. Domingo A, Marcos J. Propuesta de un indicador de la «clase social» basado en la ocupación. *Gac Sanit.* 1989; 3(10) :320-6.
219. Alonso J, Pérez P, Sáez M, Murillo C. Validez de la ocupación como indicador de la clase según la clasificación del Registro General Británico. *Gac Sanit.* 1997; 11(5):205-13.
220. Álvarez Dardet C, Alonso J, Domingo A, Regidor E. Grupo de trabajo de la Sociedad Española de Epidemiología. La medición de la clase social en ciencias de la salud. Barcelona: SG Editores; 1995.
221. Fernández E, Alonso RM, Quer A, Borrell C, Benach J, Alonso J, Gómez G. El autopercepción social como indicador de posición socioeconómica. *Gac Sanit.* 2000; 14 (3):218-25.
222. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *J Comput Graph Stat.* 1996; 5(3): 299–314.
223. Garcia Garmendia JL, Maroto Monserrat F. Interpretación de resultados estadísticos. *Med Intensiva.* 2018; 42 (6): 370-9.
224. Espejo I, Fernández F, López M, Muñoz M, Rodríguez A, Sánchez A, Valero C. Estadística descriptiva y probabilidad: teoría y problemas. Cádiz: Universidad de Cádiz; 2009.
225. Lévy Mangin JP, Varela Mallou J. Análisis multivariantes para las ciencias sociales. Madrid: Pearson-Prentice Hall; 2003.
226. Hair JF, RE Anderson, RL Tatham, WC Black. Análisis multivariante. Madrid: Prentice Hall. 1999.
227. Pérez C. Técnicas de análisis multivariante de datos. Madrid: Prentice Hall; 2004.
228. Peña D. Análisis de datos multivariantes. Madrid: Mc Graw Hill; 2002.

229. Cuadras Avellana CM. Métodos de análisis multivariante. Barcelona: EUNIBAR; 1981.
230. Alt F, Smith, N. Multivariate Process Control. En: Krishnaiah PR, Rao CR, Handbook of Statistical, Vol. 7, Amsterdam: North- Holland; 1988.
231. Bilodeau M, Brenner D. Theory of multivariate statistics. Springer; 1999.
232. Díaz Monroy LG, Morales Rivera MA. Análisis estadístico de datos multivariados. Bogotá, Colombia: Universidad Nacional de Colombia; 2012.
233. Fuchs C, Benjamini Y. Multivariate profile charts for statistical process control. *Technometrics*. 1994; 36(2): 182-95.
234. Mortimer K, Lesosky M, García-Marcos L, Asher MI, Pearce N, Ellwood E, Bissell K, El Sony A, Ellwood P, Marks GB, Martínez-Torres A, Morales E, Perez-Fernandez V, Robertson S, Rutter CE, Silverwood RJ, Strachan DP, Chiang CY; Global Asthma Network Phase I Study Group. The burden of asthma, hay fever and eczema in adults in 17 countries: GAN Phase I study. *Eur Respir J*. 2022; 60 (3): 2102865.
235. Rocco I, Cilluffo G, Ferrante G, Cibella F, Marcon A, Marchetti P, Ricci P, Minicuci N, La Grutta S, Corso B. Investigating the relationship between parental education, asthma and rhinitis in children using path analysis. *Int J Environ Res Public Health*. 2022;19 (21):14551.
236. Almqvist C, Pershagen G, Wickman M. Low socioeconomic status as a risk factor for asthma, rhinitis and sensitization at 4 years in a birth cohort. *Clin Exp Allergy*. 2005; 35(5):612-8.
237. Uphoff E, Cabieses B, Pinart M, Valdés M, Maria Antó J, Wright J. A systematic review of socioeconomic position in relation to asthma and allergic diseases. *Eur Respir J*. 2015; 46(2):364–74.
238. Escarrer-Jaume M, Juliá-Benito JC, Quevedo-Teruel S, Prieto del-Prado A, Sandoval-RuballosM, Quesada-Sequeira F, Álvaro-Lozano M. Cambios en la epidemiología y en la práctica clínica de la alergia mediada por IgE en pediatría. *An Pediatr*. 2021; 95 (1): 56. e1-56. e8.

239. D'Amato G, Pawankar R, Vitale C, Lanza M, Molino A, Stanziola A, Sanduzzi A, Vatrella A, D'Amato M. Climate change and Air pollution: effects on respiratory allergy. *Allergy Asthma Immunol Res.* 2016;8 (5):391- 5.

240. Powell P, Brunekreef B, Grigg J. How do you explain the risk of air pollution to your patients?. *Breathe.* 2016;12 (3):201-3.

241. D'Amato G, Cecchi L, D'Amato M, Liccardi G. Urban air pollution and climate change as environmental risk factors of respiratory allergy: an update. *J Investig Allergol Clin Immunol.* 2010;20 (2):95-102.

242. Galindo Domínguez H. Estadística para no estadísticos: una guía básica sobre la metodología cuantitativa de trabajos académicos. 3Ciencias; 2020.

VIII. ANEXOS

8.1. Certificado del comité de ética

El comité Ético de Investigación en Humanos de la Comisión de Ética en Investigación Experimental de la Universitat de València,

CERTIFICA:

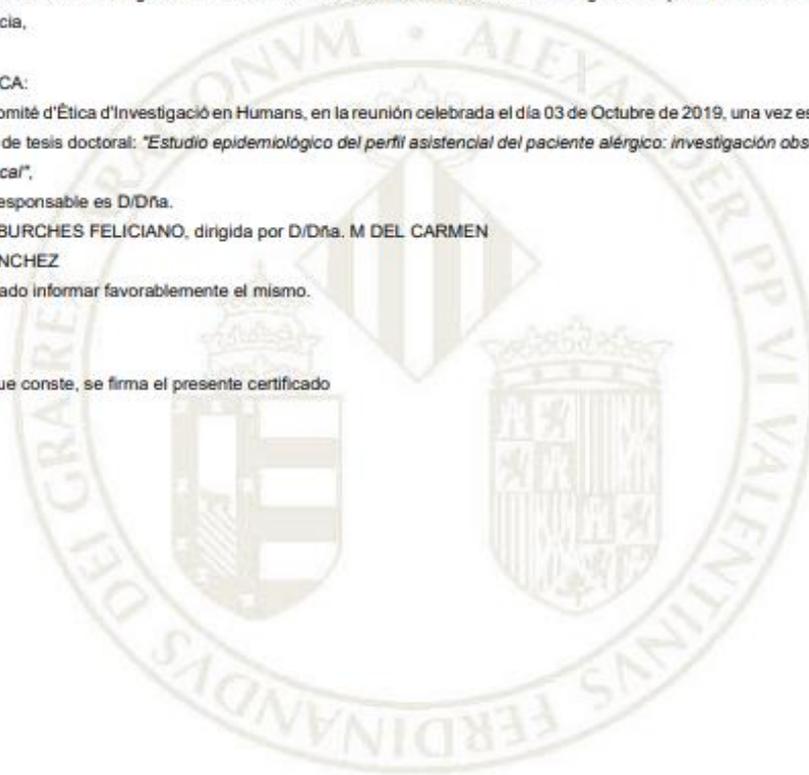
Que el Comité d'Ètica d'Investigació en Humans, en la reunió celebrada el dia 03 de Octubre de 2019, una vez estudiado el proyecto de tesis doctoral: *"Estudio epidemiológico del perfil asistencial del paciente alérgico: investigación observacional a nivel local"*,

Cuyo/a responsable es D/Dña.

MARTA BURCHES FELICIANO, dirigida por D/Dña. M DEL CARMEN
SAIZ SANCHEZ

ha acordado informar favorablemente el mismo.

Y para que conste, se firma el presente certificado



Av. Blasco Ibáñez, 13 tel: 963854109 vicerec.investigado@uv.es
València 46010 fax: 963983221 www.uv.es/serimes

Firmado digitalmente por
PEDRO JESUS PEREZ ZAFRILLA
Cargo: Presidente del Comité de Ética de la Investigación en Humanos
Fecha: 12/11/2019 10:48:05 CE

