Arch Public Health
2000, 58, 85-107

Descriptive analysis of the pharmaceutical prescription evolution in the Belgian ambulatory sector between 1986 and 1996

by

Philippart D. [1], Van Tielen R. [1], Sturbois G. [2], Peys F. [3]

Summary

The analysis of the pharmaceutical prescriptions evolution according to age groups is often restricted to some histograms. It is difficult to consider this approach if the aim is to study and compare the structural evolution of the public spending for each therapeutic class (Anatomic Therapeutic Class). The factorial analysis and classification methods allow us in this respect to see the associations between age groups and various therapeutic classes synthetically. The results obtained by classification are fully consistent with the pharmacotherapeutic and epidemiologic reality, from the standpoints of age and the pathologies underlying the prescriptions. The previously unpublished graphic representation

Facultés Universitaires Catholiques de Mons (FUCaM), 151, chaussée de Binche, 7000 Mons (Belgique), email: david.philippart@fucam.ac.be, Tel:00 32 65 32 34 24, Fax: 00 32 65 31 56 91.

Institut Belge de l'Economie de la Santé (IBES), 102, rue Léon Théodor, 1090 Bruxelles (Belgique), Tel: 00 32 2 421 15 04, Fax: 00 32 2 421 15 09.

[1] Economiste.

[2] Professeur.

[3] Pharmacien, Directeur de l'IBES.

enables decision-makers (in health or health care) to assess pharmacotherapeutic classes by age with only one graph. indeed, up to now, the sole available data concerned the type of the beneficiary (TIP (normal insured person), VIPO(widows, disabled, pensioners and orphans),...).

Key-words

Correspondence analysis, trajectories classification, evolution data, prescriptions.

Introduction

In the context of a study of the impact of demographic variables on total ambulatory pharmaceutical expenditure (1, 2), we defined a lot of tables regarding health insurance expenditure for the 44 therapeutic classes inducing the highest costs. All these data were distributed over 9 age groups and had been established based on information from IMS Health and the Belgian Pharmacists Association.

The aim of the current study is to investigate potential associations between the different age groups and the various therapeutic classes. In a second stage we integrated the time dimension in order to assess the steadiness of structures identified and to evaluate, whenever appropriate, trajectories of the therapeutic classes in terms of prescriptions number by age groups. The computer tool allows the use of more or less sophisticated statistical techniques, previously difficult to exploit. The factorial correspondence analysis was used to synthesize in a few graphs a descriptive analysis that would otherwise have been long and fastidious, or even impossible, with the traditional descriptive statistical method. This work is an unprecedented application of that method to the pharmaceutical field, particularly the section dealing with trajectories of the various therapeutic classes over time.

1. Data presentation and method

The information used below is drawn from a databank of the Belgian subsidiary of IMS Health, the main supplier of pharmaceutical statistics in the world. When confronted with series of data tables, it is crucial to have an overall view of the information base. In order to solve such a

complexity, factorial methods are probably the most adequate exploratory techniques. Both the descriptive and the evolution analyses benefited from the concomitant use of factorial and classification techniques (3, 4, 5). Theoretically, this allows taking decisions not only about the reality of defined classes but also about their relative positions, their density and their dispersion. Therefore, either technique validates the other. Our approach was to suggest a theoretical development of the part on trajectories analysis only. Factorial analysis techniques are supposed to be known but are recalled in appendix 1. As software, we used Win Spad version 3.2.

In this study we focused on data connecting three sets of objects, represented as I, J and T, in which a number noted as P_{ijt} is associated to each triplet (i, j, t) of I x J x T. This number P_{ijt} stands for the number of prescriptions for a therapeutic class i in an age group j during a period of time t. More exactly, data at our disposal relate to 44 therapeutic classes (appendix 2) and 9 age groups for periods 1986, 1990, 1993 and 1996. In other words, it is a 3-track table noted as P_{IJT}, where for each observation period (set T) prescriptions are broken down according to 44 therapeutic classes (set I) and 9 age groups (set J).

For an observation year t, the two-track table P_{IJT} presents as table 1 hereafter.

The number of prescriptions has been growing over the past few years, more particularly between 1986 and 1996. However, it should be noted that this evolution is different according to age groups. Table 2 and figure 1 illustrate clearly that increases in prescription figures are mainly due to the elderly.

TABLE 1

Visualisation of data structure for a given year (P_{IJt})

	0-4 year	5-11 year	12-19 year	20-29 year	30-39 year	40-54 year	55-64 year	65-79 year	80 year and +
therap. Cl.1									
therap. Cl.2									
⋮				*Data related to observation period t*					
therap. Cl.i				*(number of prescriptions)*					
⋮									
therap. Cl.44									

TABLE 2
Periodic variations in the number of pharmaceutical prescriptions between 1986 and 1996

	Number of prescriptions (000)	Periodic growth	Growth since 1986
1986	44 824		
1990	47 784	6.6% (1.6%)	6.6% (1.6%)
1993	56 003	15.7% (5.0%)	24.9% (3.2%)
1996	61 352	9.6% (3.1%)	36.9% (3.2%)

NB: The values in brackets represent the mean annual variations for the considered period.

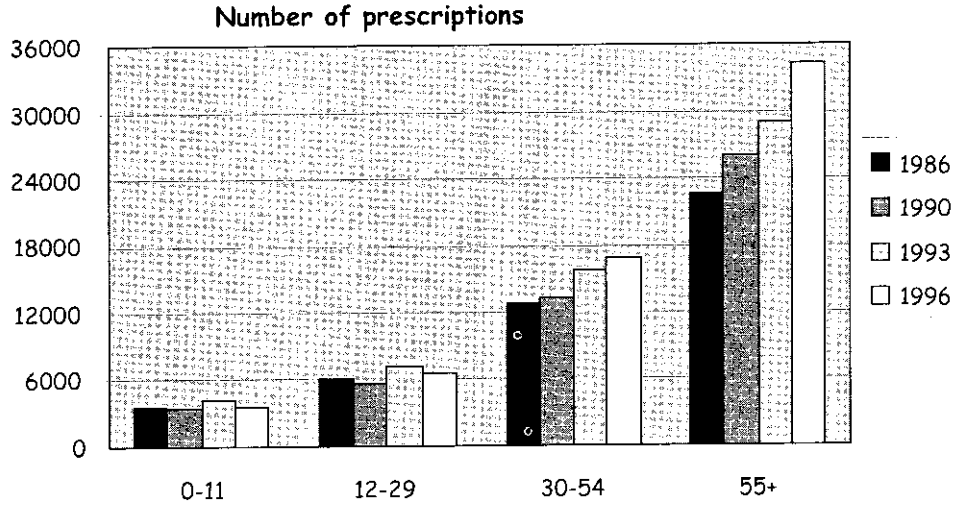


Fig. 1: Evolution of the number of ambulatory pharmaceutical prescriptions by age groups (in thousands)

2. Analysis of associations between therapeutic classes – age groups

The above statements made it possible to establish the context of the study and to roughly describe available data. At this stage we studied the associations that might exist between the different age groups and the various therapeutic classes based on the number of prescriptions.

We have a three-track table at our disposal, represented as P_{IJT}, where for each observation period (set T), prescriptions are broken down respectively according to therapeutic classes (set I) and age groups (set J).

This table allows us now to investigate associations between therapeutic classes and age groups. We first performed a factorial correspondence analysis of table $P_{IJ, 96}$ (prescriptions for the year 1996). This analysis is complemented with a classification of therapeutic classes based on their profiles in terms of prescription numbers by age groups.

2.1. Correspondence analysis (table $P_{IJ, 96}$)

The plane defined by the first two factor axes presents 63.33% + 22.26% = 85.59% of the total inertia. This plane contains most of the available information.

To properly interpret this graph, it should be remembered that if prescriptions of one therapeutic class distribute exactly on the same pattern as the total number of prescriptions for the combined age groups, this therapeutic class will be positioned exactly at the origin of the axes (6).

In the present case, the trial shows a progression of age groups through the pharmaceutical "map". Axis 1 opposes on the one hand specialties more specifically aimed at children and teenagers and on the other hand therapeutic classes typical of the elderly. Furthermore, the age groups arranged themselves in a parabolic form known as the Gutman effect, so that the second axis merely qualifies the interpretation of the first factorial axis.

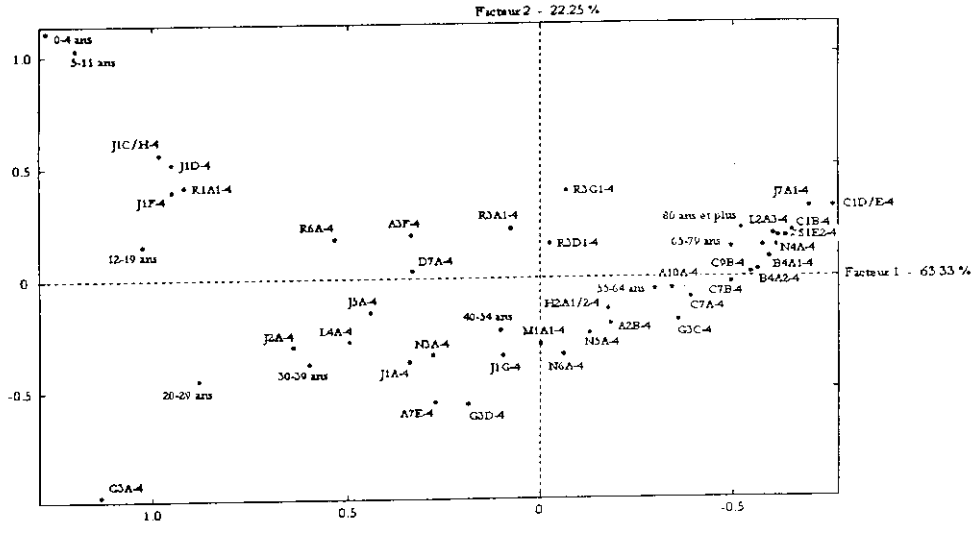


Fig. 2: Representation of the therapeutic classes profiles for the 1996 (table $P_{IJ, 96}$)

If the relative weight (in terms of prescription figures) of a therapeutic class in the first age groups is above average and gradually regresses for the intermediate age groups to finally be under average in the last age groups (J1D (cephalosporin antibiotics), R1A1 (nasal corticoïdes), J1F (macrolide antibiotics), J1C/H (penicillines), the position of this therapeutic class on the graph will be linked to that of the left portion of the parabola.

If the relative weight of a therapeutic class in the first age groups is below average and gradually increases for the subsequent age groups, the position of this therapeutic class on the graph (C8A/B, calcium antagonists, used in cardiovascular diseases for instance) will be linked to that of the right portion of the parabola and means a gain of its relative weight on the age scale.

If the relative weight of a therapeutic class is below average in the first and last age groups and is higher in the intermediate age groups ($A7^e$, gastrointestinal antiinfectives for instance), this specialty will be positioned at the top of the parabola.

If the relative weight of a therapeutic class is higher in the first and last age groups and lower in the intermediate ones (R3G1, combined antiasthmatic drugs for instance), it will be located inside the parabola.

From a strictly pharmacotherapeutic standpoint (2, 7, 8), figure 4 shows a strong epidemiological coherence. As an example, let us define the position of a few main therapeutic classes in terms of prescriptions. Let us start with the most expensive class i.e. class A2B (gastrointestinal antiulcer drugs). We see that the prescription weight is located at the level of age groups 40-54 years and 55-64 years; this can easily be explained since this class treats a chronic condition typical of people stressed by active life and having bad eating habits. As a consequence, prescriptions for class B4A (lipid-lowering drugs) are usually located in the immediately next age group: these drugs are mainly prescribed for the prevention of cardiovascular diseases (resulting among others from elevated cholesterol levels). This might reflect an awareness of the high number of deaths of cardiovascular origin. The launch and the reimbursement (though conditional (9)) of new molecules led to a substantial increase in the number of prescriptions. With regard to NxAs, i.e. drugs administered in mental illnesses, we can note how logical the position of the different therapeutic classes in the age groups is. N3A drugs (antiepileptic drugs) are mainly prescribed for younger adults (< 40 years), although clinical signs may appear at all ages. The position of N5A drugs

(antipsychotics) is somewhat more peculiar. And indeed, the prescription weight lies relatively more to the right than expected. Psychogeriatrics, a highly progressing discipline, might explain this position. It should not be overseen that schizophrenia and paranoia can also occur earlier in life. Antidepressants (N6A) also hold an unexpected position. Older people (55 years and over) seem to focus a great deal of prescriptions, with the result that the use of antidepressants is not explicitly linked to younger age groups. The positioning of antiparkinson drugs (N5A) is logical, since Parkinson's disease is typically a condition of old age. Furthermore they are positioned at the same level as the therapeutic class treating eye conditions such as glaucoma (S1E2), a disease also typical for the very old age. Pulmonary conditions, treated by the various therapeutic classes R3Ax, including intrinsic and/or allergic asthma, do not seem to be linked to age but rather to environmental factors. In this case prescriptions do not focus on a specific age group. Conversely, nasal allergies are treated by the R1A1 class and expenditure is restricted to the teenage group (hay fever, ...).

With regard to antibiotics, i.e. the J1X classes, we can confirm that each age group has its own range of drugs. So J1A (tetracyclines, for chronic bronchitis) and J1G (quinolones, for urinary infections) classes mainly concern the adults and the elderly while other classes, such as penicillins (J1Cs) and macrolides (J1Fs), are usually prescribed in pediatric acute infections. Prescriptions of influenza vaccines (J7A) reflect perfectly well the willingness of our health authorities to protect patients over 60, who enjoy reimbursement (40%). And indeed, various Belgian and foreign studies have demonstrated the cost/effectiveness or benefit of this strategy in terms of lives saved and public spending reduced.

2.2. Classification of therapeutic classes (table $P_{IJ, 96}$)

In order to complete the graphical representation of the therapeutic classes profiles, a classification was carried out from factor scores on the first three axes (63.33% + 22.25% + 9.42% = 95%). To this end we used the Ward method (hierarchical ascending classification), which allowed us to raise the following dendrogram.

This representation is a step towards choosing one or several partitions of the combined therapeutic classes. The review of the tree form should lead to the definition of the most appropriate partitions. The best partition seems to divide the hierarchical tree into six classes.

From a technical standpoint, it should be said that dividing a tree into partitions involves interlocking of partitions. The definition of a partition can

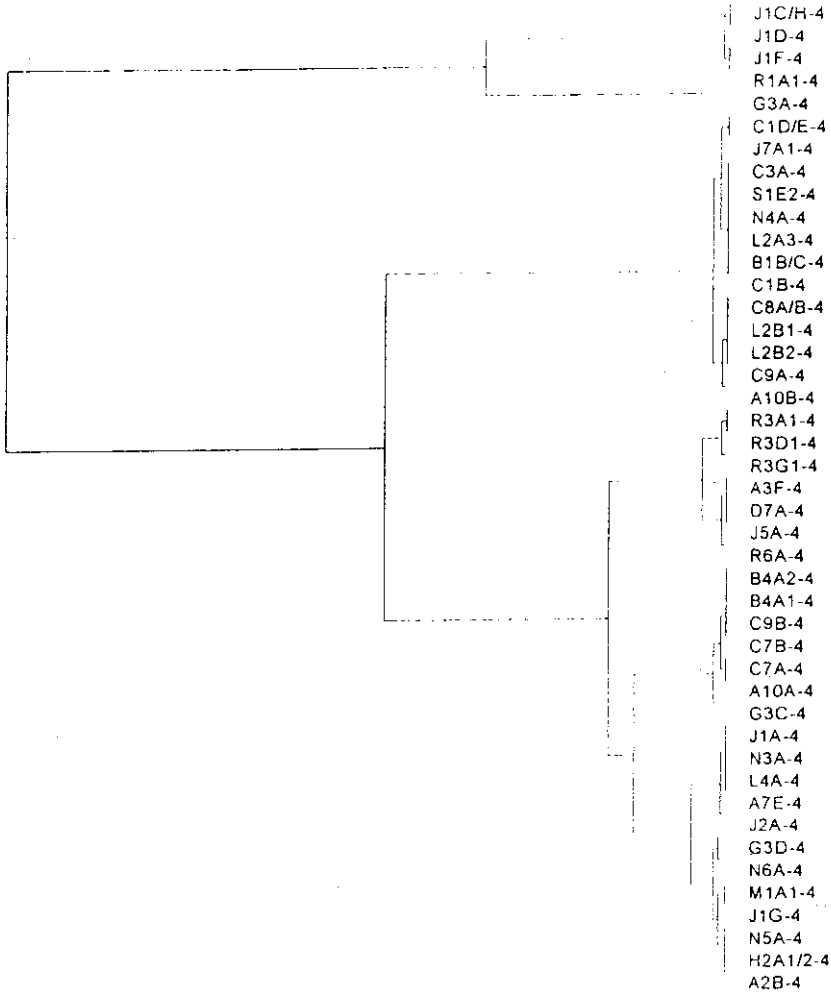


Fig. 3: Tree representation resulting from the classification of pharmaceuticals (table $P_{IJ,\ 96}$)

thus be improved by relieving this constraint. Therefore, to improve the cluster homogeneity, we carried out a partition "consolidation": it consists of a few iterations with mobile centers, these centers being initially the centers of the classes obtained by dividing the tree.

These clusters are easy to identify from a factorial standpoint, but are they appropriate from a pharmacotherapeutic standpoint?

Groups 2, 4, 5, 6 have a relatively obvious justification. Group 2 represents prescriptions issued prophylactically for the old age group. The main conditions treated are hypertension and hyperlipidemia. Group 4 includes prescriptions for therapeutic classes aimed at treating typically

TABLE 3

Clustering of therapeutic classes based on their 1996 profiles (table $P_{IJ, 96}$)

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
A2B J1G N6A	A10A C9B	A3F R3D1	A10B C8A/B L2B2	G3A	J1C/H
A7E L4A	B4A1 G3C	D7A R3G1	B1B/C C9A N4A		J1D
G3D M1A1	B4A2	J2A R6A	C1B J7A1 S1E2		J1F
H2A1 N3A	C7A	J5A	C1D/E L2A3		R1A1
J1A N5A	C7B	R3A1	C3A L2B1		

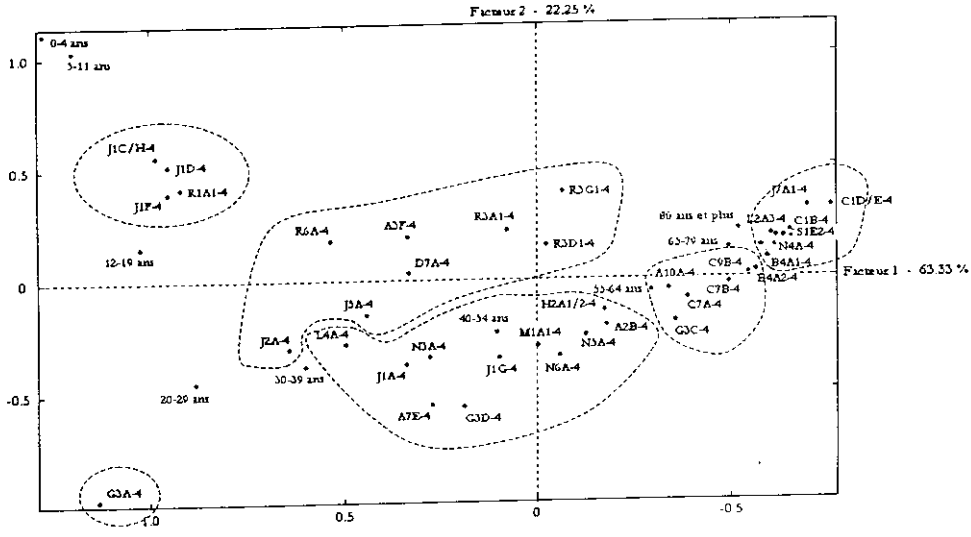


Fig. 4: Representation of the groups of therapeutic classes defined based on their 1996 profiles (table $P_{IJ, 96}$)

geriatric conditions with obvious clinical signs (Parkinson's disease, glaucoma, coronary heart disease). Group 5 does not require many explanations since it relates to female oral contraception with an undoubtedly temporal specificity. Group 6 represents the traditional pediatric treatments (infections).

The interpretation of groups 1 and 3 repartition appears less obvious. For group 1, the relative weight (number of prescriptions) is below the average number of prescriptions (taken as a whole) in the first and last age groups and is higher in the intermediate age groups i.e. the active population age groups. This means that all the therapeutic classes concerned have a trough position on the parabola. We find in this group therapeutic classes with the highest number of prescriptions, on the one hand antidepressants, antiepileptics, neuroleptics, i.e. drugs treating the majority of mental illnesses, and on the other hand therapeutic classes

treating conditions typical of the modern world, such as anti-inflammatory drugs which are not only prescribed for rheumatism but also for sport or professional conditions such as back pain, ...

Group 3 is positioned more clearly inside the parabola, which means that the therapeutic classes concerned are prescribed either for younger adults (R6As or antihistaminics against hay fever, J2As or antimycotics against vaginal candidosis), or for older adults (R3X drugs treating respiratory tract diseases such as asthma, COPD, ...). The intermediate age groups are less involved. We can thus ascertain the epidemiological coherence of this computerized classification.

However, it would be interesting to characterize these clusters by age groups. For instance, to support the statement previously issued concerning group 6 being positioned left on the factorial plane (J1D, R1A1, J1F, J1C/H): if the relative weight of the number of prescriptions for a pharmaceutical is above average in the first age groups and gradually decreases in the intermediate age groups to finally get below average in the last age groups, the graphical situation of this therapeutic class will be linked to that of the left portion of the parabola.

Before going on with the analysis, one could question the steadiness of the structure observed in 1996 regarding the progression of age groups

TABLE 4

Characterization of the groups of therapeutic classes defined on their 1996 (table $P_{IJ, 96}$)

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
0-4 year	☺	☺	☹	☺	☺	☹
5-11 year	☺	☺	☹	☺	☺	☹
12-19 year	☺	☺	☹	☺	☹	☹
20-29 year	☹	☺	☹	☺	☹	☹
30-39 year	☹	☹	☹	☺	☹	☹
40-54 year	☹	☹	☺	☺	😐	☺
55-64 year	☹	☹	☺	☹	☺	☺
65-79 year	☺	☺	☺	☹	☺	☺
80 year and more	☺	☺	☺	☹	☺	☺

😐: Relative weigth (according to prescription number) of the considered group is equivalent to the mean relative weigth (all groups disconcerted).

☺: Relative weigth (according to prescription number) of the considered group is above the mean relative weigth.

☹: Relative weigth (according to prescription number) of the considered group is lower the mean relative weigth.

through the "map" of pharmaceutical classes. In order to answer this question, a validity analysis using re-sampling methods was performed; it allowed us to validate the positioning of age groups as a parabola.

Furthermore, one can verify that this structure is steady over time, i.e. there is no rotation of axes over time, percentages of variance explained remain in the same range and age groups describe a similar progression in the factorial plane for the 4 observation years. The correspondence analysis performed on the first three tables separately ($P_{IJ, 86}$, $P_{IJ, 90}$, $P_{IJ, 93}$) allows us to study trajectories described by therapeutic classes against this structure.

3. Analysis of trajectories described by therapeutic classes

In view of the structure identified above, it would be interesting to study the progress of therapeutic classes against that structure, what we visualize as trajectories (4, 10). In order to assess this progress for each therapeutic class, we took data over from 1990, 1993 and 1996, and we projected the therapeutic classes as additional items on the factorial plane F1F2 obtained from the correspondence analysis that had been carried out on subtable $P_{IJ, 86}$ (1986 data). Four points are thus associated to each therapeutic class. By joining these four points, one can have an idea of the progress of each therapeutic class separately in terms of prescription numbers. As a result, each therapeutic class has its own trajectory in the factorial plane.

For instance, therapeutic classes A3F and N4A display rather different trajectories [1].

The two pharmaceuticals differ very clearly in their average position in the factorial plane, in the sense that N4A is more strongly connected to the right portion of the parabola, and in their trajectories in this plane. The N4A prescription repartition among age groups is steady over time, while the A3F one undergoes an abrupt change between 1986 and 1990 before getting steady again.

[1] The darkest point corresponds to time t = 1986 and the lightest one to time t = 1996.

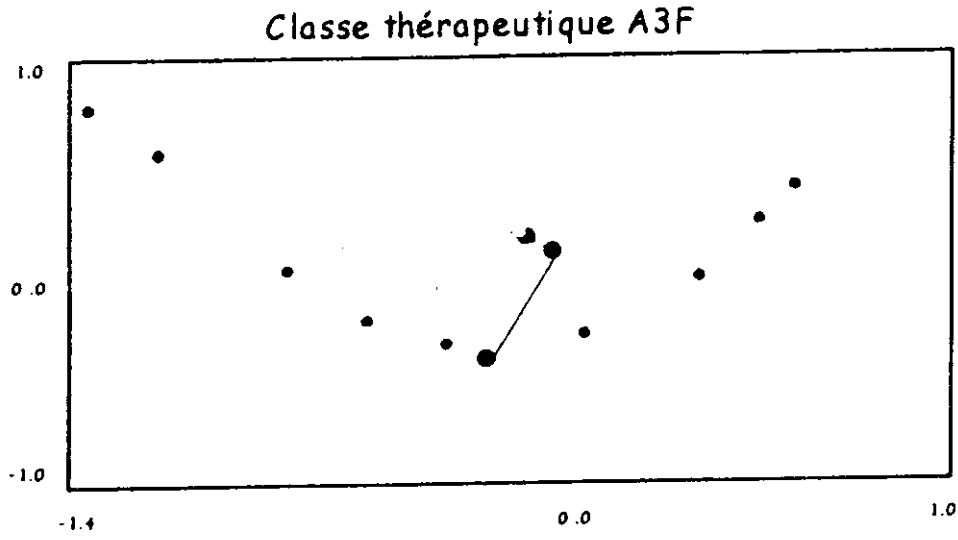


Fig. 5: Trajectories of therapeutic classes A3F in the factorial plane F1-F2

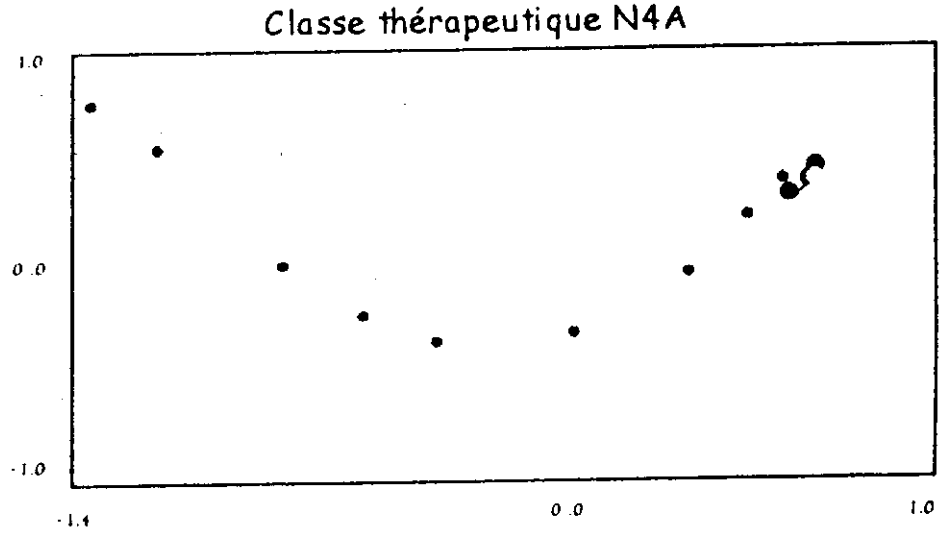


Fig. 6: Trajectories of therapeutic classes N4A in the factorial plane F1-F2

What we are more interested in is the classification of trajectories described by therapeutic classes. Indeed, the visualization of trajectories in the factorial plane is difficult to interpret due to the high number of trajectories. A trajectory classification would allow a better overview by regrouping the therapeutic classes with similarities in their structure progression over time. This would enable to reduce the number of trajectories by assimilating the trajectories of the therapeutic classes of each group with the trajectory of the center of gravity of the group.

TABLE 5
Clustering of trajectories described by therapeutic classes based on their progress (table P_{ijt})

Group 1						Group 2	Group 3	Group 4
A2B	C1B	C9A	J1C/H	M1A1	R3D1	A3F	J7A1	L2B2
A7E	C1D/E	D7A	J1F	N3A	R6A	J1D		
A10A	C3A	G3C	J1G	N4A	S1E2	J5A		
A10B	C7A	G3D	J2A	N5A		L2B1		
B1B/C	C7B	H2A1/2	L2A3	N6A		R1A1		
B4A2	C8A/B	J1A	L4A	R3A1		R3G1		

The first group, made up of 32 therapeutic classes, is characterized by a strong steadiness of the relative weight of prescriptions by age group during the period considered (1986-1996). This steadiness is consistent with our intuition, owing to the epidemiology in general and to the physician prescription habits in particular, which remains stable over time. Group 2 includes therapeutic classes treating conditions with a growing incidence such as antivirals (aids), respiratory affections (asthma), some cancer types. Factors linked to the market can also interfere (introduction of new products, decline of others). In other words, all therapeutic classes don't grow old in the same way.

One group is made of a single therapeutic class (L2B2); its atypical behavior can be explained by a structural measure, i.e. the introduction of a conditional reimbursement for some types of cancer. Enclosed are the detailed trajectories of the main therapeutic classes.

Discussion and conclusion

The representation of the therapeutic classes profiles (figure 4) provides a condensed visualization of phenomena that would otherwise require many of the more traditional graphs and tables. And indeed, one can identify at a glance therapeutic classes inducing prescriptions more specifically linked to some age groups. This original approach allows an instant overview of age groups with specific conditions or granted specific public health measures.

Overall, this representation of therapeutic classes profiles reflects the outstanding therapeutic coverage of our population, irrespective of age groups. Besides, it summarizes many conclusions of studies conducted on pharmaceutical prescription and consumption according to age. From

table 3 we know that people aged 55 and over account for 56% of the total number of pharmaceutical prescriptions (in 1996) while their demographic weight is only 26.5%. This phenomenon, remarkably explained among others by Mizrahi and Mizrahi, would be the result of the number of diseases experienced by the elderly. This observation is strongly highlighted on figure 4 with a high concentration of therapeutic classes in the area of people over 65. In order not to make a specific group of the population feel guilty, it should be specified that Mizrahi and Mizrahi have also demonstrated that for the same number of diseases, the pharmaceutical consumption is identical irrespective of age groups. The progress of the number of prescriptions might also be explained by the medical staff and our health authorities becoming aware of the importance to adequately prevent a certain number of diseases. In an aging population the aim is to keep a maximum of patients healthy. Thus, pharmaceuticals can no longer be limited to curative treatments.

The review of ambulatory prescriptions in the various therapeutic classes by age groups has revealed a highly steady structure through the periods studied. The representation of the prescription progress by means of factorial trajectories makes it possible to integrate the time component and, owing to its synthetic aspect, it is also a very convenient way to visualize changes in the structure of each therapeutic class. Furthermore, a classification can easily be defined from trajectories described by therapeutic classes based on their progress. This approach casts specific light on the reality without providing directly usable information due to the steadiness of the relative number of prescriptions by age groups and by therapeutic classes. These results are fully consistent with the health professionals' intuitions.

Résumé

L'analyse évolutive des prescriptions pharmaceutiques selon les tranches d'âge se limite souvent à quelques représentations sous forme d'histogramme. Cette approche est difficilement envisageable si l'on désire étudier et comparer l'évolution structurelle des dépenses pour chaque classe thérapeutique (ATC). Les méthodes d'analyse factorielle et de classification permettent à cet égard de visualiser de manière synthétique les associations existant entre les tranches d'âge et les diverses classes thérapeutiques. Les résultats obtenus par la classification sont tout à fait cohérents avec la réalité pharmaco-thérapeutique et épidémiologique tant au point de vue de l'âge que des pathologies sous-jacentes à la prescription. La représentation graphique inédite permet aux décideurs politiques (de la santé ou des soins de santé) d'évaluer les classes pharmaco-thérapeutiques par tranche d'âge avec un seul graphique. En effet, jusqu'à présent, les seules données disponibles concernaient le type de bénéficiaire (Titulaire Indemnisable Primaire, VIPO,...).

Mots-clés

Analyse des correspondances, classification des trajectoires, données évolutives, prescriptions.

References

1. VAN TIELEN R, PEYS F, GENAERT J. The demographic impact on ambulatory pharmaceutical expenditure in Belgium. *Health Policy* 1998; 45: 1-14.
2. PEYS F, VAN TIELEN R. Les dépenses INAMI pour les spécialités pharmaceutiques délivrées en milieu ambulatoire: analyse rétrospective des grandes classes thérapeutiques. *Annales Pharmaceutiques Belges*, Août-Septembre 1998, 29-32.
3. ESCOFIER B, PAGÈS J. Analyses factorielles simples et multiples: objectifs, méthodes et interprétation. Dunod, 1990.
4. JAMBU M. Méthodes de base de l'analyse des données. Eyrolles, 1999.
5. LAVIT CH. Analyse conjointe de tableaux quantitatifs. Masson, 1988.
6. LEBART L, MORINEAU A, PIRON M. Statistique exploratoire multidimensionnelle. Dunod, 1995.
7. ACCORSI B. La consommation pharmaceutique et ses motifs, CREDES, Mai 1995, n° 1074, série thèses et mémoires, 108 pages.
8. MIZRAHI AN, MIZRAHI AR. Evolution de la consommation pharmaceutique par personne et par maladie. Credes, mai 1995, n° 1071.
9. PEYS F, VAN TIELEN R. Le remboursement conditionnel des médicaments en Belgique. *Annales Pharmaceutiques Belges,* Mars 1999, 23-26.
10. DAZI F, LE BARZIC JF. L'analyse des données évolutives: méthodes et applications. Technip, 1996.

Appendix 1: Presentation of the statistical analysis methodology

In the context of the present study, several statistical analysis methods are considered, among which correspondence analysis and classification.

Correspondence analysis

In view of the descriptive and multivariate nature of the chosen approach, factorial correspondence analysis should apply well to this problem and to the available data. As observed by Bouroche and Saporta, with this exploratory method, it is possible to study similarities between individuals with suggestive graphical representations.

This statistical method is based initially on the study of contingency tables. A contingency table contains the frequency of association between the modalities of two qualitative variables. Generally speaking, correspondence analysis is warranted when, at the intersection of a row and a column, a table contains a value that can be assimilated to a sample and that characterises the row and the column. Such a table can be written:

$$K_{I,J} = \{K(i,j) | i \in I, j \in J\}$$

Where

I: all modalities of a nominal variable
J: all modalities of another nominal variable
K(i, j) or k_{ij}: value that characterises the couple (i, j)

Such a table is a privileged field of application for correspondence analysis, by comparing the profile of rows and columns (i.e. conditional relative frequencies) based on the chi-squared index, with a view to showing the associations between variables. More precisely, deviations towards independence are divided into a number of factors by this analysis. Therefore, this technique analyses the dependence structure between qualitative variables and shows its main characteristics.

As for the choice of the distances, it may be useful to justify the use of the chi-squared. The usual Euclidean distance between two lines of the raw k_{IJ} table would only represent the differences in samples between two modalities. On the other hand, the Euclidean distance between two

profiles/lines is a good representation of the similarity or the difference between two modalities in terms of distribution. Yet, this distance favours columns, with their important mass. In addition, it does not check the property of distribution equivalence, while the latter is fundamental, as it guarantees a certain amount of result invariance towards the nomenclature chosen for constructing the modalities of a variable, provided similar modalities are grouped. For all these reasons, each deviation is weighted by the inverse mass of the column. A new distance, called $\square^2$ distance, is calculated:

$$d^2(i,i') = \sum_{j=1}^{p} \frac{1}{f_j}\left(\frac{k_{ij}}{k_i} = \frac{k_{i'j}}{k_{i'}}\right)^2 \quad \text{where}$$

$$k_i = \sum_{j}^{p} k_{ij}, f_j = \sum_{i}^{n} \frac{k_{ij}}{k} \quad \text{and}$$

$$k = \sum_{i}^{n}\sum_{j}^{p} k_{ij}$$

Graphically, this data analysis technique produces a representation of the modalities of sets I and J by points that form a scatter diagram in a multidimensional space. The centre of the scatter diagram is called the center of gravity. This subspace has a system of orthonormal axes that cross the centre of gravity and hence give the "best possible" representation of the scatter diagram form. (Note that in a correspondence analysis, the scatter diagram is a subspace with p-1 dimensions, with p being the smallest dimension in the table).

These axes can be organised into a hierarchy based on the inertia index. The first one has the highest inertia, i.e. the variance of the points projected onto this axis is maximal. For each axis, it is possible to calculate a percentage of inertia, i.e. the part of the total inertia on the axis in question. (The total inertia of the scatter diagram measures the distance of the points towards their center of gravity, i.e. the global dispersion of the diagram). This measure reflects the quality of the representation of the axis. One of the interests of a factorial analysis is to keep only those axes that are relevant, i.e. those that give the best synthesis of the information contained in the initial data table. Only factors with an inertia index above a threshold conventionally defined as the ratio between total inertia (100%) and the smallest dimension of the table are kept.

The data that do not fit with the independence hypothesis reflect associations between rows and columns perceived through proximity.

Therefore, two modalities are more related when their representative points are closer to one another and further away from the origin. Similarly, points corresponding to opposed profiles (that vary in opposite directions) will generally be exceptionally far away from one another.

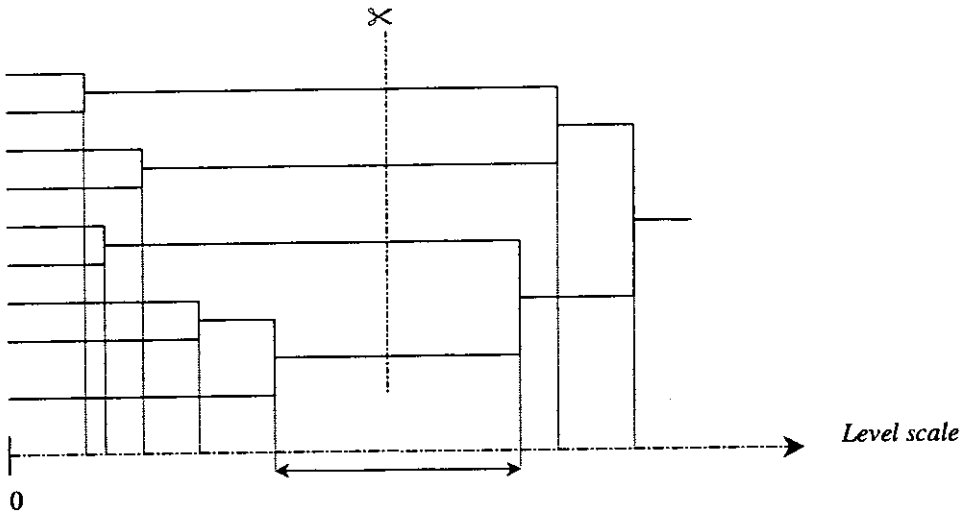
For each axis derived from correspondence analysis, it is possible to calculate absolute and relative contributions of the elements introduced in the analysis. The relative contribution measures the quality of the representation of each point: the closer this coefficient to the unit, the closer the point to the axis. Hence, it is a measure of the proximity of points to axes. Absolute contributions determine which points are most characteristic of an axis through a measurement of the contributions of the different modalities to the inertia of the axes.

Classification

Factorial analysis defines the mutual positioning of the objects to be studied on graphs and provides continuous spatial representations. However, usually, this analysis is completed by classification, as factorial methods do not take account easily of higher-order interactions. In addition, certain structures may be so complex that projections in subspaces are not sufficient. Therefore, a typology is often a handy observation method, beyond the first dimensions provided by a factorial analysis.

The objective of classification or nomenclature methods is to group objects in order to define a number of homogeneous classes as far as their profile is concerned. A typology is obtained when each group of objects is converted into an entity with known characteristics. In the context of our analyses, we used an ascending hierarchical classification method.

A hierarchical tree or dendrogram is built using Ward's criterion. Ward's criterion is based on the minimal reduction of variance through pooling and compatible with the inertia criterion used for the determination of factorial axes. The tree is a step in the choice of one or several divisions in the object set. By building the tree and studying its form, it is possible to gain a certain insight into the likely number of classes, as dense zones of homogeneous points are clustered at the lowest level of the tree and separate into classes as linkages get longer. The hierarchy produced in this way is indexed: each segregation corresponds to a numerical value that represents a pooling level. The higher the index, the more heterogeneous the pooled parts.



Hierarchical tree: depending on the form of the tree and the level scale, it seems that the most likely segregation is the one that generates four classes: as we progress towards the summit (i.e. from left to right), the most significant deviation increase between two levels occurs between the third and fourth values of the index.

The analysis of the tree leads to the choice of the segregation(s) that look(s) most probable. Then, the tree should be "cut" to create the desired segregation(s).

Finally, it should be noted that, in the present case, the classification algorithm is based on the factorial coordinates produced by a preliminary correspondence analysis. This will allow refining the initial configurations and, therefore, obtaining more adequately typed categories by abandoning the last factorial axes, which often include random components.

Appendix 2

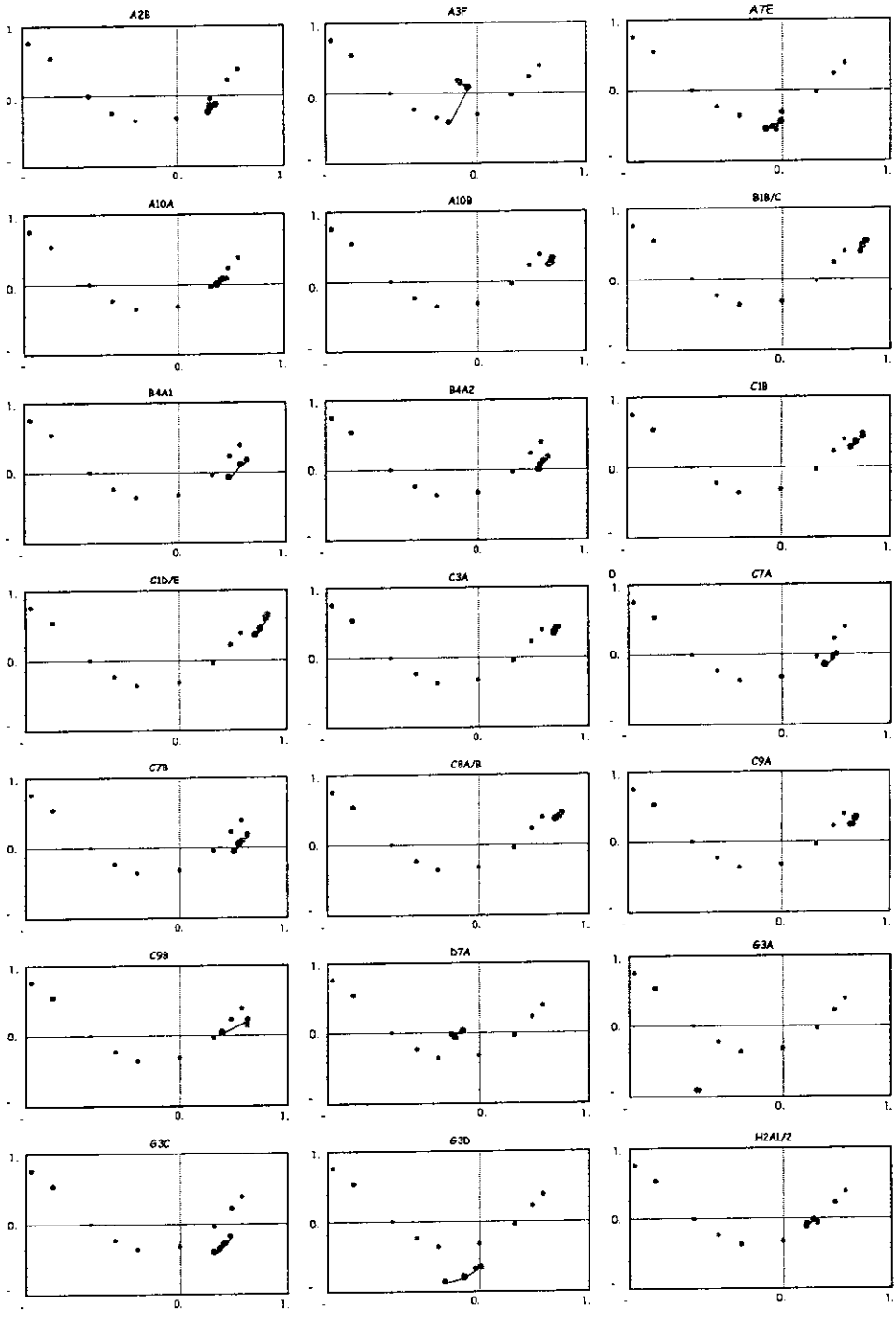
Therapeutic classes

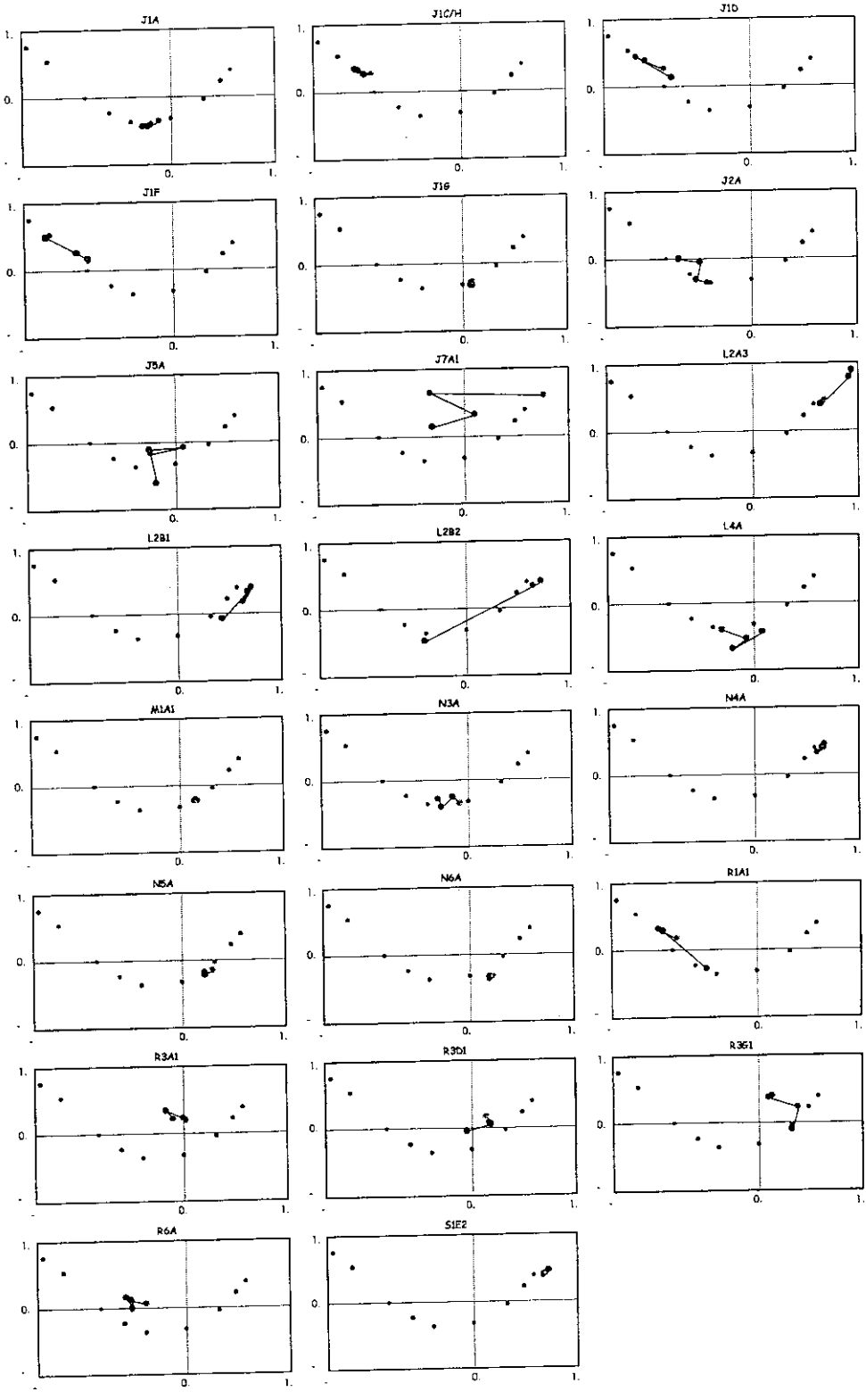
A2B	Antiulcerants
A3F	Gastroprokinetics
A7[e]	Intestinal Anti-Inflammatory Agents
A10A	Insulins
A10B	Oral Anti-Diabetics
B1B/C	Anticoagulants Injectable and Platalet Aggregation Inhibitors
B4A1	HMG-CoA Reductase Inhibitors (Statins)
B4A2	Fibrates
C1B	Anti-arrhythmics
C1D/E	Coronary Therapy and Nitrites and Nitrates
C3A	Diuretics
C7A	Beta-Blocking Agents Plain
C7B	Beta-Blocking Agents Combinations
C8A/B	Calcium Antagonists Plain and Combinations
C9A	ACE Inhibitors Plain
C9B	ACE Inhibitors Combinations
D7A	Topical Cortoicosteroids
G3A	Hormonal Contraceptives, systemic
G3C	Oestrogens and Combinations
G3D	Progestogens and ombinations
H2A1/2	Systemic Corticosteroids, Plain (injectable and oral)
J1A	Tetracyclines and Combinations (Antibiotics)
J1C/H	Broad Spectrum Penicillins (Antibiotics)
J1D	Cephalosporins (Antibiotics)
J1F	Macrolides and similar types (Antibiotics)
J1G	Fluoroquinolones (Antibiotics)
J2A	Systemic Agents for Fungal Infections
J5A	Antivirals excluding Vaccines
J7A1	Vaccines Grippe (Influenza)
L2A3	Cytostatic Gonadotrophin-releasing Hormones Analogues
L2B1	Cytostatic anti-oestrogen
L2B2	Cytostatic anti-androgens
L4A	Immunosuppressive Agents
M1A1	Antirheumatics non-steroidal Plain
N3A	Anti-Epileptics
N4A	Anti-Parkinson drugs
N5A	Psycholeptics
N6A	Anti-depressants / Thymoanaleptics

R1A1	Nasal Corticosteroids Without Anti-Infectives
R3A1	B2-Stimulants Inhalants
R3D1	Corticoids Inhalants
R3G1	Anticholinergics-Plain and Combinations with B2-stimulants
R6A	Systemic Antihistamines
S1E2	Miotics and antiglauma Preparations – Topical

The prescriptions of these 44 therapeutic classes stand for more than 84% Belgian public spending of pharmaceuticals delivered in ambulantory sector.

Appendix 3: Representation of the therapeutic classes trajectories




J1A
J1C/H
J1D
J1F
J1G
J2A
J5A
J7A1
L2A3
L2B1
L2B2
L4A
M1A1
N3A
N4A
N5A
N6A
R1A1
R3A1
R3D1
R3G1
R6A
S1E2