Article

Costs and mortality associated with HIV: a machine learning analysis of the French national health insurance database

Martin Prodel,[1] Laurent Finkielsztejn,[2] Laëtitia Roustand,[3] Gaëlle Nachbaur,[3] Lucie de Leotoing,[1] Marie Genreau,[1] Fabrice Bonnet,[4,5] Jade Ghosn[6,7]

[1]HEVA, Lyon; [2]ViiV Healthcare, Rueil-Malmaison; [3]GlaxoSmithKline, Rueil-Malmaison; [4]CHU de Bordeaux, Service de Médecine Interne et Maladies Infectieuses, Hôpital Saint-André, Bordeaux; [5]Université de Bordeaux, INSERM U1219, ISPED, Bordeaux; [6]Assistance Publique – Hôpitaux de Paris, APHP. Nord-Université de Paris, Hôpital Bichat-Claude-Bernard, Service des Maladies Infectieuses et Tropicales, Paris; [7]INSERM U1137, IAME, Université de Paris, France

Abstract

Background: The objective is to characterise the economic burden to the healthcare system of people living with HIV (PLWHIV) in France and to help decision makers in identifying risk factors associated with high-cost and high mortality profiles.

Design and methods: The study is a retrospective analysis of PLWHIV identified in the French National Health Insurance database (SNDS). All PLWHIV present in the database in 2013 were identified. All healthcare resource consumption from 2008 to 2015 inclusive was documented and costed (for 2013 to 2015) from the perspective of public health insurance. High-cost and high mortality patient profiles were identified by a machine learning algorithm.

Results: In 2013, 96,423 PLWHIV were identified in the SNDS database, including 3,373 incident cases. Overall, 3,224 PLWHIV died during the three-year follow-up period (mean annual mortality rate: 1.1%). The mean annual *per capita* cost incurred by PLWHIV was € 14,223, corresponding to a total management cost of HIV of € 1,370 million in 2013. The largest contribution came from the cost of antiretroviral medication (M€ 870; 63%) followed by hospitalisation (M€ 154; 11%). The costs incurred in the year preceding death were considerably higher. Four specific patient profiles were identified for under/over-expressing these costs, suggesting ways to reduce them.

Conclusions: Even though current therapeutic regimens provide excellent virological control in most patients, PLWHIV have excess mortality. Other factors such as comorbidities, lifestyle factors and screening for cancer and cardiovascular disease, need to be targeted in order to lower the mortality and cost associated with HIV infection.

Key points

This study provides an estimate of the total economic burden of HIV infections to public health insurance in France. This disease cost remains high (€ 1,370 million annually), and we suggest ways to reduce it through a study of associated factors, as well as associated mortality. The identification of high-cost patient profiles is enabled by a machine learning algorithm.

Introduction

In 2013, the number of people living with HIV (PLWHIV) in France was estimated at around 153,400,[1] of whom 24,500 were unaware of their HIV status and a further 13,800 had been diagnosed but were not yet receiving an antiretroviral treatment (ART).[2] In 2013, 20,126 PLWHIV were hospitalised at least once for a reason related to their HIV infection.[3] Given that combined ART (cART) is lifelong, that complications of HIV infection or cART are frequent and that PLWHIV are at increased risk for certain comorbidities, the economic burden of HIV infection is currently high and may increase in the future.

Nonetheless, the extent of this burden remains poorly characterised. A review of the literature on economic studies of the management of PLWHIV in five Western European countries performed in 2014 highlighted the paucity of the available data and the disparity of the findings.[4] Several reasons were put forward to account for this, including small sample sizes[5] and incomplete cost assessment. Most of recent economic studies were modelling studies, principally evaluating cost-effectiveness of new ARTs,[6] assessing the impact of early diagnosis of high-risk groups,[7,8] or of on-demand pre-exposure prophylaxis.[9] The only cost-of-illness study identified, which collected real-world data, measured hospital costs only,[3] but there are little data to date assessing the global burden of HIV infection.[10-12]

A consolidated national healthcare database (SNDS) is available in France for health economic studies. It includes exhaustive healthcare resource consumption data, covering both hospital and community medicine sectors, for all individuals with public National Health Insurance (NHI) in France.[13] The SNDS is

Significance for public health

The authors believe that this study is important in that it provides for the first time an estimate of the most impacting factors of the total economic burden of HIV infections to public health insurance in France. The approach relies on a state of-the-art machine learning analysis. Regarding epidemiological results, although the prevalence and incidence of HIV infections in Western European countries are reasonably well documented, there is a paucity of information on its economic burden. We suggest ways that the economic burden of HIV (€ 1,400 million annually to public health insurance), as well as associated mortality, could be reduced. The main interest and novelty of the results from a public health perspective is the identification of patient profiles at risk of high costs. Such results provide clear and potentially actionable levers to decision makers, either for prevention plans, for patient pathways, or for reimbursement strategies.

exhaustive for the NHI and was proven reliable to conduct epidemiological studies.[13] The use of machine learning technics to predict clinical outcomes from such electronic health records or claim data provides is an ever-increasing topic of interest for decision makers.[14-20]

The main objectives of this study were to estimate the annual cost of management of PLWHIV at a national level in France. Costs are reimbursed care from the NHI's perspective. The corollary research question was to determine if patient characteristics were linked to lower/higher management costs and mortality, and to find such profiles.

Design and methods

Study design

This study is a retrospective analysis of PLWHIV identified in the French NHI database (SNDS: Système National des Données de Santé) initiated in 2018. This database covers all healthcare consumption by French residents insured by the general NHI regimen which cover around 75% of the French population and is representative of the entire population in age and gender.[13] Data from other insurance schemes are also available in the SNDS, but not exhaustively collected, especially regarding mortality. The study identified all PLWHIV recorded in the database in 2013. All healthcare consumption for these individuals was tracked backwards since 2008 to identify comorbidities and forwards to 2015. 2008 is the earliest year for which fully linked information on healthcare consumption is available. The cost analysis was performed from the perspective of the NHI. Mortality was assessed on the three-year period 2013-2015 and was compared with standardised mortality rates in the French general population[1].

Study population

The study population consisted of all PLWHIV identified in the database by at least one of the following criteria: i) eligibility for long-term disease status (ALD) due to HIV infection in 2013, ii) at least two prescriptions for a specific anti-HIV treatment in 2013-2014, iii) HIV gene sequencing for drug resistance testing or iv) hospitalisation with a mention of HIV infection on the discharge summary. Patients that had not been affiliated with the same general insurance fund since 2008 were also excluded as data from other insurances are unavailable. Although it lowers the cohort size, it avoids falsely underestimating the primary outcome of this study, costs of care, due to long-lasting unrecorded care in patient's claim history.

The prevalent population was defined as all eligible subjects in 2013, and an incident population as those among them with no HIV-related event recorded before 2013. Individuals were excluded if no healthcare resource consumption was recorded between 2013 and 2015 if they were <18 years and had been treated with an ART for less than seven weeks (ART for six weeks is a recommended prophylaxis in children born to HIV-positive mothers).

Data extraction

Data were extracted for each eligible patient concerning age, gender, date of death (if deceased) and any eligibility for ALD status. All healthcare resource consumption (hospitalisations, consultations, medical procedures and tests, and medication) and reimbursement (date and amount) over the period between 2008 and 2015 inclusive was documented.

The comorbidities were identified based on either chronic disease identified in hospital discharge summaries, ALD status or long-term medication prescription for vascular prevention or psychiatric disorders. Pathologies of particular interest in PLWHIV (hepatitis B or C infection, non-Hodgkin lymphoma and Kaposi's sarcoma), as well as opportunistic infections potentially attributable to HIV infection, were documented on the basis of ICD-10 codes in hospital discharge summaries for the period 2008–2015 or with ALD status for one of these conditions (Supplementary Table 1).

Costing

Costs were assigned from the French national tariffs applicable in the year of resource consumption, and were expressed in euros, updated to 2019-values to take into account inflation. Two costs were measured: the annual cost per incident patient, from the inclusion date, and the overall reimbursed costs for the calendar year 2013. For hospital stays, a standard national tariff was applied based on DRG codes. These standard tariffs include medical and related procedures, nursing care, treatments (except specific expensive drugs), food and accommodation.

For the study of factors associated with the annual cost, the focus was to find subgroups of patients (= profiles) under-expressing or overexpressing such cost. To strengthen the detection of the most prominent profiles, and to provide clear and explainable results to decision makers, three arbitrary categories (low, medium, high) were created based on the observed cost distribution. Thus, the search for patient profiles was addressed as a 3-class classification problem.

Statistical methods - decision tree analysis

A machine learning model using decision tree analysis was built to study the factors associated with the risk of having low/high annual cost of care, using a binary splitting decision tree algorithm (scikit-learn library 0.21.2, python 3.7).[19,21] The model built itself iteratively, starting with the entire cohort (the root, the centre of the sunburst) and first searched for the most discriminating variable to divide the cohort into two subgroups (the branches, the innermost circle) which segregated as much as possible the target variable (the annual cost) into three arbitrary categories (low, medium, high). The resulting tree is represented as a "sunburst plot".

All variables collected from the SNDS database were considered as potential discriminating variables and entered in the algorithm. Fifty-eight individual variables were considered, including demographics, socioeconomic variables, comorbidities, pathologies related to HIV infection and healthcare resource consumption variables, all listed in the Supplementary Table 2. The quality of segregation was estimated by the Gini impurity (the lower the Gini, the most discriminating a variable is).[21] In each following round, each branch was split into two further sub-branches using the same procedure (on the sunburst plot, each segment of the circle is divided in two sub-segments). Further sub-branches were developed until one of three criteria were met, namely when the resulting subgroup corresponded to <1% of the original cohort, when the decrease in Gini impurity was minimal (<0.1) or when six levels of division from the root had been reached. These criteria prevent the algorithm from overfitting the data and ensure concise identified profiles.

The model was initially built on a random sample of 80% of the study cohort. Its performance was then tested on the remaining 20% of the cohort. A decision tree technic was chosen, among other machine learning technics, for its capability to perform well on multiclass classification, and for the purpose of clarity and interpretability of the results.[14,15,22]

Then, the association of study variables with mortality (binary

outcome) was first evaluated using multivariate logistic regression, performed with regularisation to adjust the impact estimation of relevant features. Associations were expressed as logit (log odds-ratio) functions. Then, a machine learning model was used to find high-risk profiles, similarly to costs. The two technics complement each other: the logistic regression emphasizes and compares the intensity of risk protector factors, while the decision tree provides actionable profiles for decision makers.

Results

Incident and prevalent cases of PLWHIV in 2013

Overall, 126,515 PLWHIV fulfilling the eligibility criteria for the study were identified in the SNDS database in 2013. After exclusion of patients who had not been affiliated with the same general insurance fund since 2008, 96,423 individuals were available and constituted the prevalence population. For 3,373 of these, no HIV-related event was identified before 2013, and these constituted the incident population (Figure 1). In the incident population, 521 individuals (15.5%) were at CDC C stage at the time the HIV infection was diagnosed.

Overall, 93.4% of the prevalent population had been prescribed an ART at least once, in most cases a triple cART. The mean age was 47.5 (±12.0) in the prevalent population and 39.4 (±14.2) in the incident population, with the age distribution shifting slightly towards older ages in men (48.7±12.0) compared to women (45.1±12.0). Sociodemographic and clinical features of the prevalent population are presented in Supplementary Table 3.

Comorbidities

Overall, 61.9% of the prevalent population were eligible for ALD status because of a comorbidity other than HIV. The most frequent of these were chronic liver disease, psychiatric disorders and 'cardiac and neurovascular diseases' (Supplementary Table 3). Hepatitis B infections were identified in 7.0% of individuals and

hepatitis C infections in 12.5%. Opportunistic infections were identified in 3.4% of individuals.

Cost analysis

The total cost of management of the 96,423 prevalent PLWHIV was € 1,370 million in 2013, of which the largest contribution came from the cost of ART (Table 1). The distribution of total annual *per capita* costs of management of HIV between 2013 and 2015 is presented in Figure 2. The mean annual *per capita* cost was € 14,223±€ 10,800 and the median cost € 12,093 [interquartile range: € 9,000 to € 16,795]. *Per capita* costs tended to increase with age, although a plateau was reached around the ages of 40 to 50 (Supplementary Figure 1).

Table 1. Total cost of care for HIV by type of expenditure.

| Expenditure type	Total cost (M€)
TOTAL	€ 1,370
Antiretroviral therapy	€ 869.8 (63.4%)
Hospitalisations	€ 154.3 (11.3%)
Sick-leave benefit	€ 116.3 (8.5%)
Any other medication	€ 73.2 (5.3%)
Outpatient visits	€ 40.7 (3.0%)
Paramedical support	€ 20.6 (1.5%)
Specialist consultations	€ 19.2 (1.4%)
Medical devices	€ 18.2 (1.3%)
Transportation costs	€ 18.1 (1.3%)
Laboratory tests	€ 16.0 (1.2%)
General practitioner consultations	€ 14.5 (1.1%)
Other outpatient care	€ 9.6 (0.7%)
Other HIV-specific medication	€ 7.1 (0.5%)
Medical care provided at the patient's home	€ 1.0 (0.1%)
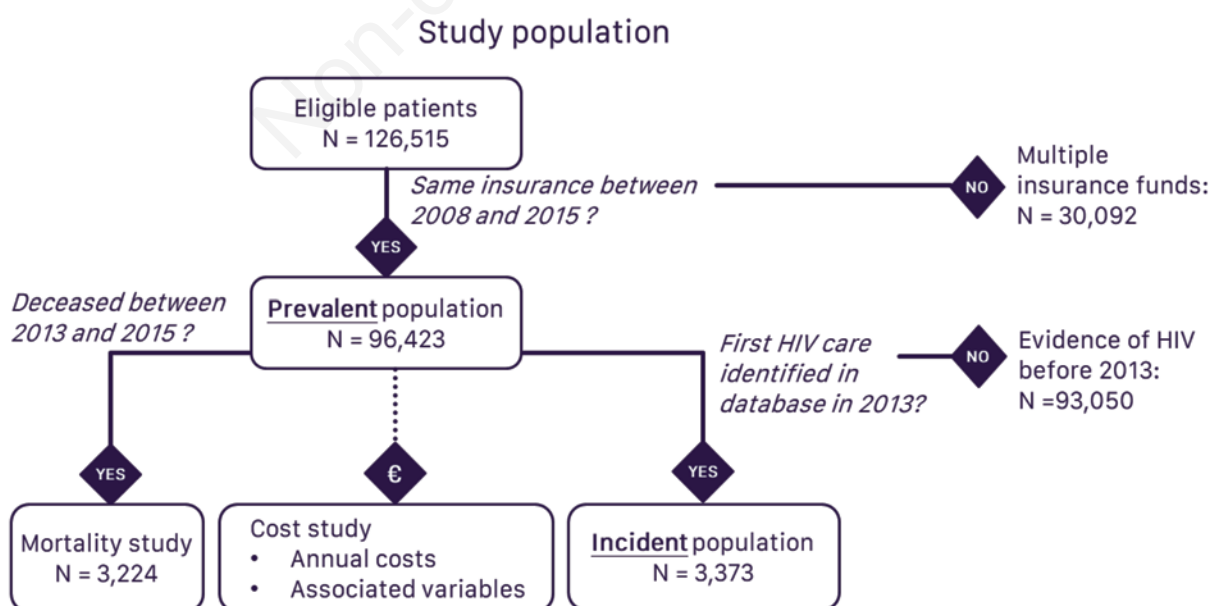
Figure 1. Study population.

For the patients who died during the follow-up period (2013-2015), the costs incurred in the year before they died were considerably higher, with a mean *per capita* annual cost of € 45,189±€ 36,314 (median: € 35,951) in men and € 48,505±€ 37,625 (median: € 41,754) in women.

Variables associated with cost

To focus on extreme costs, and based on the cost distribution shown on Figure 2, we identified the 20% and 80% percentiles. To improve the clarity of results toward public health decision makers, the thresholds were set to the closest rounded thousand. It results in three categories, namely high (≥€ 18,000), moderate (€ 8,000 – € 18,000) or low (<€ 8,000), which were validated by the scientific committee. These thresholds are well suited to the cost statistical distribution (right skew) and avoid emphasising outliers.

Globally, the two principal drivers of cost identified in the decision-tree analysis were the number of months without any documented healthcare reimbursement (accounting for 43% of variation in costs) and hospitalisation (29% of variation) (Table 2). Since healthcare resource consumption expressed in volume is expected to covary tightly with cost, it was decided to reiterate this analysis excluding these two variables from the model input (dropout strategy). This reduction enables the detection of differences among patients without care gaps nor hospitalizations. In this reiteration of the model, the four principal cost drivers were the number of comorbidities (34% of variation in cost), the duration of HIV treatment before inclusion (20%), the first class of cART prescribed (14%) and the number of non-AIDS defining infections identified during hospitalisation (14%) (Table 2).

Further information about four outstanding profiles resulting from this analysis is presented in Figure 3. Profile 1, accounting for 10% of the cohort, was associated with high cost (≥€18,000) and characterised by the presence of two or more associated comorbidities, an absence of non-AIDS defining infections and multiple ART treatment. Profile 2, also associated with high cost, accounted for 4% of the cohort and was characterised by the presence of two or more associated comorbidities and at least one non-AIDS defining infection. Profile 3 was associated with moderate cost (€8,000 to €18,000) and accounted for 61% of the cohort. This cohort was characterised by no more than one comorbidity and having a standard ART regimen throughout the follow-up. Profile 4, associated with low cost (<€8,000), accounted for 5% of the cohort and was characterised by no more than one comorbidity and no recorded ART treatment.

Mortality

Overall, 3,224 patients of the prevalent population died during the follow-up period (1,015 in 2013, 1,144 in 2014 and 1,065 in 2015), corresponding to a mean annual mortality rate of 1.1%. Age and gender specific mortality rates in the prevalent population were compared with standardised mortality rates in the French general population. Mortality rates in HIV individuals from the SNDS were higher than in the general population in all age groups (for 2013: 40-44 age group: 6.13 vs 1.38; 45-49: 9.75 vs 2.29; 50-54: 12.66 vs. 3.62). The excess mortality in PLWHIV tended to decline over the age of sixty in men (for 2013: 65-69 age group: 19.48 vs 15.10), whilst it increased in women (for 2013: 65-69 age group: 17.75 vs 6.60).

Variables associated with mortality

In the prevalent population, the available variables presenting the strongest associations with mortality in the multivariate logistic regression analysis were hospitalisation, older age and active cancer (predisposing factors). General practitioner consultations, the number of non-AIDS defining infections and periods of ≥6 months
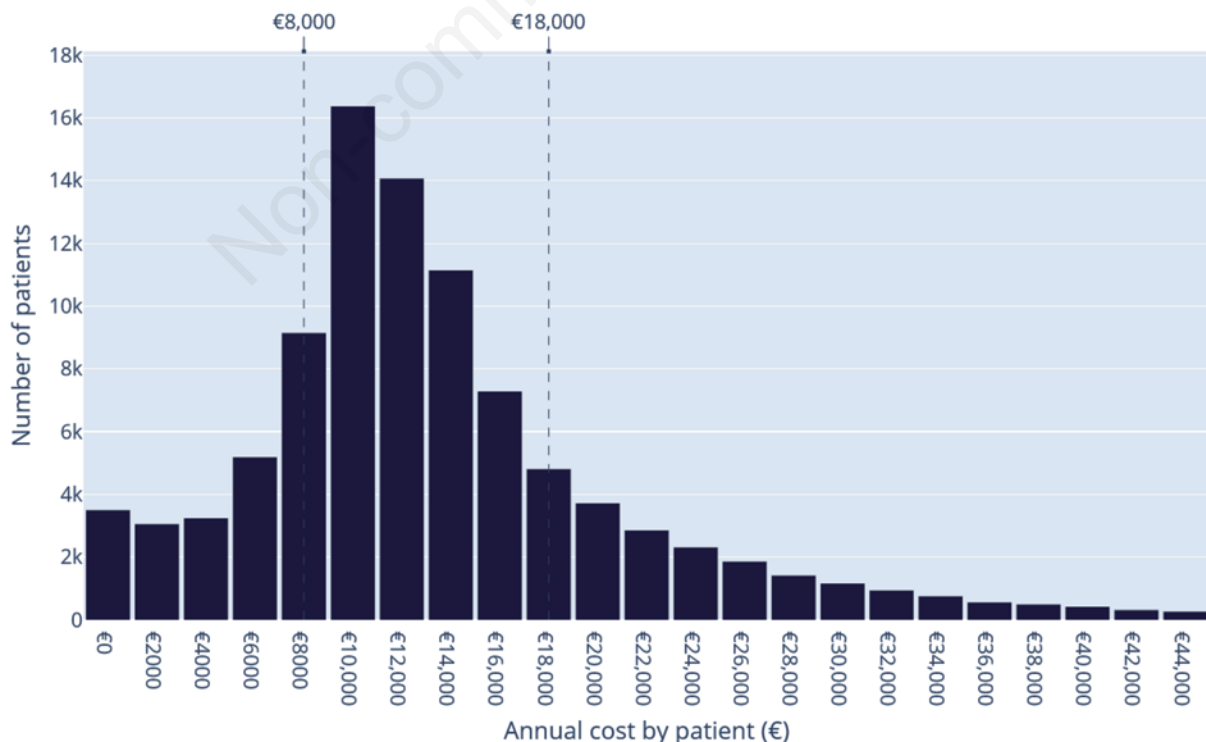

Figure 2. Distribution of total annual per capita costs of management of HIV (n=96,423).

without healthcare reimbursement were associated with decreased mortality (Figure 4A). For this regression analysis, the area under the curve (AUC) of the receiver operating characteristics (ROC) curve was 0.93.

In the decision tree model, the variables which explained the greatest amount of the variance were hospitalisation in the year before death, GP consultations and in-hospital resuscitation (Figure 4B). The AUC was 0.93. The decision tree model was reit-

Table 2. Variables contributing to the variance in cost, evaluated from 2013 to 2015. The direction of variations due to variables' contribution is shown in the profiles of Figure 3.

Population evaluated Variables excluded	MODEL 1 Entire cohort None	MODEL 2 Entire cohort Months without HCC, Hospitalisations
Variables		
Number of months without HCC	43%	-
Hospitalisation (presence/absence)	29%	-
Number of comorbidities	12%	34%
Treatment duration coverage (in months)	10%	20%
Number of infections	3%	14%
Type of first prescribed treatment (mono/bi/triple/ quadruple therapies)	-	14%
Number of ED visits	-	5%
6-month gap in HCC (presence)	-	4%
Booster (pharmacokinetic enhancers) as part of first-line therapy	2%	-
Number of treatment changes	<1%	-
Age at inclusion	<1%	-
Number of switches between triple therapies	<1%	4%
Year HIV diagnosed	<1%	<1%
Day hospitalisation	<1%	-

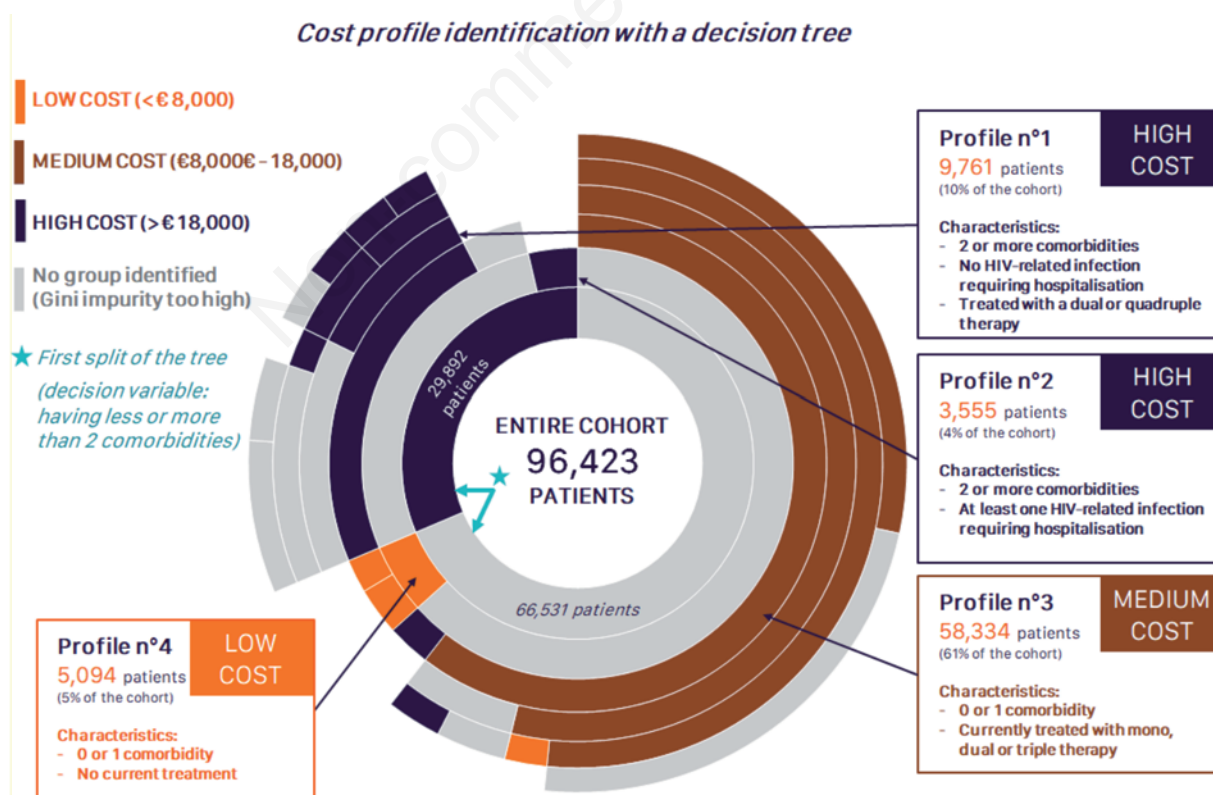ED, Emergency Department; HCC, HealthCare Consumption; HIV, human immunodeficiency virus.

Figure 3. Sunburst plot of variables associated with costs of HIV: 4 profiles identified by a machine learning algorithm.

erated after the exclusion of certain variables with an obvious relationship with mortality, namely hospitalisation, in-hospital resuscitation, age and 'no documented health care reimbursement for ≥6 months'. In this model, the five variables explaining the most vari-

ance were the number of GP consultations (48% of variance explained), the number of emergency department visits leading to hospitalisation (29%), the number of comorbidities (9%), active cancer (7%) and day hospitalisation (3%). The AUC was 0.90.
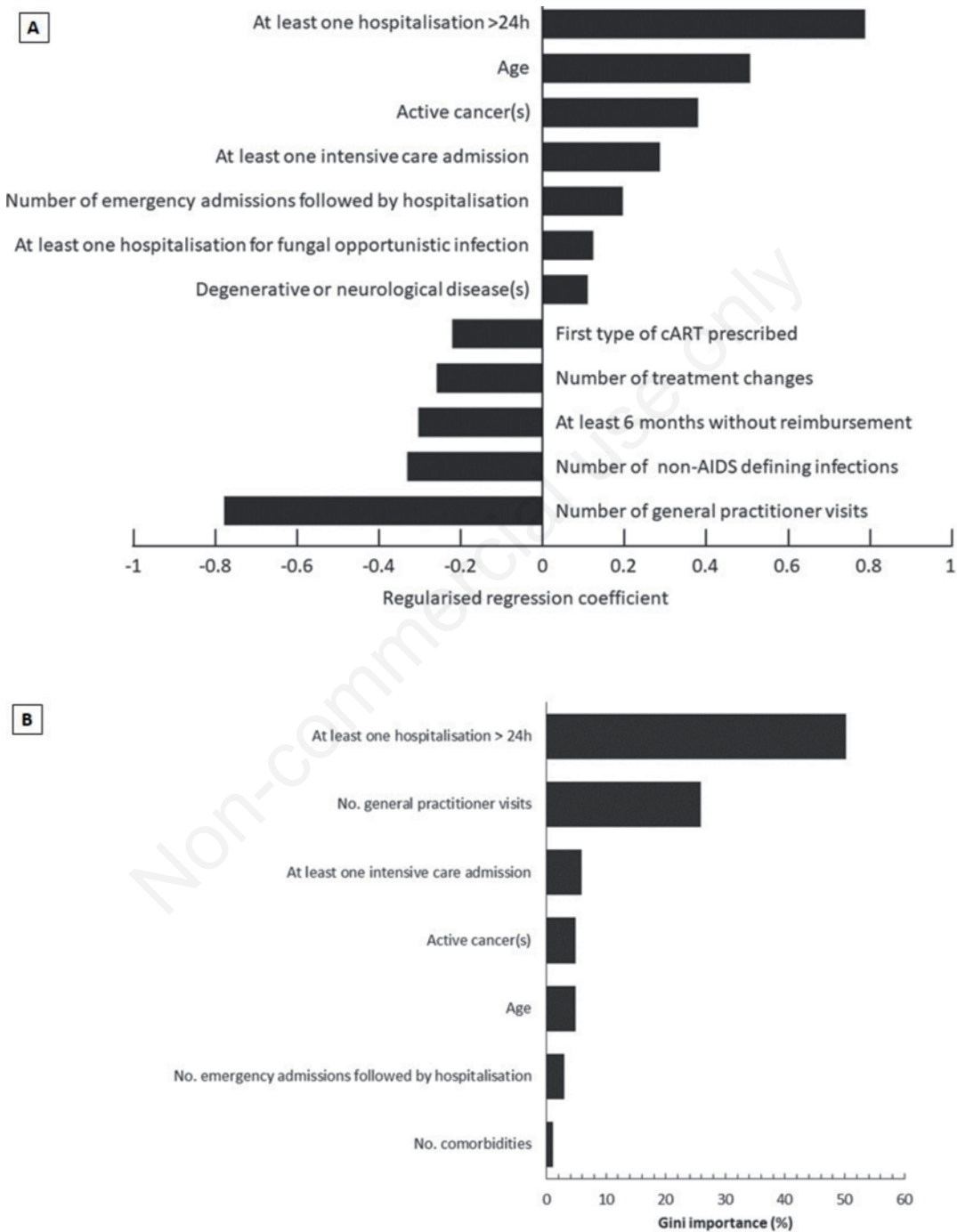

Figure 4. Variables associated with mortality (top: 4A, regression analysis; bottom: 4B, decision tree analysis). First type of cART prescribed refers to the number of associated ARTs (e.g., dual therapy, triple therapy, etc.). Result: both models, although different by nature (parametric regression and non-parametric decision tree), highlight the strong individual roles of the same variables. Variable combinations are found in profiles. [1]*Source from the French National Institute of Statistics and Economic Studies: https://www.insee.fr/fr/statistiques/2554599?sommaire=1912926*

Discussion

Study cohort

This study identified PLWHIV from the SNDS database, which covers all health resource consumption by beneficiaries of the French general NHI scheme. During the year 2013, a 'footprint' of HIV infection could be identified for 126,515 individuals in the database. This figure can be compared with the estimate of 128,586 individuals with a known HIV infection in the total French population in the estimate made previously for the same year using only the criterion of eligibility for long-term disability status due to an HIV infection.[1] The number of cases can also be compared with the number of individuals hospitalised with an HIV infection in 2013, which was 20,216.[3] This indicates that, at least for France, most PLWHIV are not hospitalised every year and that estimations of the burden of HIV based solely on hospital data will grossly underestimate the total cost of HIV to NHI.

The number of incident cases was estimated at 3,373 individuals, which is lower than the 6,600 new diagnoses of HIV identified each year.[2] This is explained by the restriction to patients at the general NHI regimen (75% of the population), and the filter on patients with at least one HIV care management, between 2013 and 2015. In addition, around 10% of HIV-seropositive individuals in France are not being cared for within the health system.[1] A limitation of this study, especially regarding the identification of at-risk profiles, is to exclude patients changing from insurance regimen. A future study could focus on this specific population.

Costs

The mean annual *per capita* cost incurred by PLWHIV was € 14,278, corresponding to a total cost of management of HIV of € 1,370 million. This can be compared, for example, to the total cost to the French NHI in 2016 of € 3,440 million for chronic respiratory diseases and of € 8,120 million for diabetes.[23] Corresponding mean annual *per capita* costs were € 929 for chronic respiratory diseases and € 2,150 for diabetes.[23] Such comparisons indicate that although the prevalence of HIV infection (~4/1000 in the French adult population in 2013[1]) is relatively low compared to other medical conditions, its management is relatively costly, especially the year prior to death. In comparison, the mean annual *per capita* cost was $ 38,439 in the USA for the same period;[11] € 19,103 in Germany from 2006 to 2009; $ 17,352 in Canada for 2017. The findings of the present study can be compared with those of a micro-costing study performed in an HIV care reference centre in the North of France in 2004-2005,[25] which reported average yearly costs of care of PLWHIV ranging from € 19,240 to € 36,540 according to CD4 count. These costs are much higher than those reported here. This difference may potentially be explained firstly by a shift from the in-hospital to the outpatient setting for routine monitoring of PLWHIV and secondly to the improved tolerability of ART which now engenders less costs related to management of side-effects and monitoring than was previously the case.[10]

The principal contributor to the cost incurred by PLWHIV was the cost of ART (61.3% of total costs). This outweighs by far the cost of hospitalisations, (12.3% of total costs). This finding is consistent with previous suggestions that evaluations of the economic burden of HIV infection based solely on hospital-based healthcare resource consumption will largely underestimate the total burden.[4] Since this was not a case-control study, it was not possible to evaluate how much of the cost was specifically related to the management of HIV.[11] In the case of ART, 100% of the cost is supposedly HIV-related. In the case of hospitalisation, a previous analysis of the national hospital discharge database in 2013 identified the total cost of specific HIV-related hospital management to be € 64 million,[3] suggesting that around 35% of hospital costs incurred by PLWHIV are specifically related to the management of HIV, with the remainder being related to management of comorbidities. Nevertheless, the cost of ART is expected to decrease in the future. This will be a consequence firstly of the wider availability of generic ARTs[6] and secondly of simplification of therapeutic regimens through the use of effective and well-tolerated dual therapies or intermittent triple therapies.[24] A reduced incidence of HIV infections can also be expected due to more widespread pre-exposure prophylaxis.

The evaluation of factors associated with the cost of care of PLWHIV principally identified factors that are sources of cost, such as hospitalisation and continuous healthcare consumption, as well as the number of comorbidities and the number of non-AIDS defining infection leading to hospitalisation. It should, however, be hypothesised that patients interrupting their treatment, and thus generate short-term economies to the healthcare system, might generate higher costs later on, when they return to the health system after a period of uncontrolled viraemia with consequences on immunity and/or HIV resistance. None of the sociodemographic variables included into the models were strongly predictive of cost. It might suggest that the overall quality of care of management of PLWHIV in France is relatively uniform and that there are no large segments of the PLWHIV population which generate a disproportionate amount of costs, as intended by national health policies. A limitation of the presented machine learning work is to address a 3-class classification problem, for clarity and communication purposes toward public health decision makers. Future works could include a sensitivity analysis on the classification thresholds, as well as performance comparisons with regressions methods.

Mortality

Around 1,100 deaths of PLWHIV per year were identified. This estimate is consistent with the estimate from a national cohort study performed in 2010 of around 1,000 deaths.[26] Even though >90% of the study population was taking ART, mortality in PLWHIV was slightly higher than in the total population of NHI beneficiaries in the SNDS database for all age groups. This excess mortality appeared somewhat higher in women, particularly in the >65-year-old age groups. However, this probably reflects the lower mortality rate for women compared to men in the general population; the absolute mortality rates in men and women with HIV were similar. In an analysis performed in France in 2000 addressing the causes of mortality in PLWHIV who were well-controlled by ART,[27] the principal causes of death were HIV-unrelated cancers, notably lung and liver cancer, cardiovascular disease and HCV-related liver disease. High rates (>35%) of HCV infections, intravenous drug use and alcohol dependence were observed in these patients. In the present study, the global rate of HCV infection was 12.5%, reflecting the declining HCV seroprevalence in France (currently <1%),[28] and mortality due to this cause may be expected to decrease.

Concerning variables associated with higher risk of mortality, identified in the multiple logistic regression analysis and the decision tree analysis, these were essentially unsurprising and included overnight hospitalisation, emergency department visits or in-hospital resuscitation and comorbidities, notably cancers. The variables most strongly associated with reduced mortality were the number of visits to a general practitioner, the number of non-AIDS defining infections and the presence of gaps in reimbursement of health-

care longer than six months. Visits to a general practitioner is a difficult variable to interpret, but may be a surrogate marker for health awareness and care-seeking behaviour. The number of documented non-AIDS defining infections may also be a consequence of care-seeking behaviours. Gaps in healthcare reimbursement may be a surrogate marker for patients who feel well and do not see any benefit to take a treatment or had adverse effects. Importantly, none of the socioeconomic variables, such as markers of social deprivation, were identified as being associated with increased mortality.

Conclusions

The principal strengths of this cost-of-illness study were that the database covers >75% of the French population and includes exhaustive information on all reimbursements for healthcare made by the NHI fund. For this reason, the derived cost information can be considered regarding the real total cost of healthcare incurred by PLWHIV in France. A light limitation is that the HIV infection (defined by serological testing or viraemia) is not explicitly documented in the database, and had to be derived from proxy variables (ART prescription, screening for ART resistance, eligibility for ALD status based on an HIV infection, or hospitalisation with a mention of HIV infection on the discharge summary). Likewise, individuals with undiagnosed HIV infections, who account for around 16% of the total population of PLWHIV,[1] cannot be identified, and cost incurred by these individuals cannot be determined. Regarding the estimation of costs, the database contains no information on important biological variables such as virological control and immunological markers which are the major determinants of long-term overall health and prognosis in PLWHIV. Data from the French national HIV registry (FHDH ANRS CO4) indicates that around 90% of patients receiving ART in 2013 had undetectable viraemia.[1] Finally, certain older hospitalised patients may be discharged to a nursing home, where healthcare consumption and deaths are not yet well-linked to the SNDS database.

In conclusion, we have shown in this analysis at a national level that mortality of PLWHIV remains higher than in the general population despite effective antiretroviral therapy. The direct medical costs of management of PLWHIV to public NHI in France are, on a *per capita* basis, relatively high compared to other chronic morbidities. Even though current ART regimens provide excellent viraemic control in most patients, the cost of therapy is high and PLWHIV still have excess mortality and HIV-related hospitalisations. Factors other than virological control, such as reducing the burden of comorbidities, addressing lifestyle factors such as smoking, and screening for cancers and cardiovascular disease, need to be targeted to lower the mortality and cost associated with HIV infection.

Correspondence: Martin Prodel, HEVA, 186 Avenue Thiers, 69006 Lyon, France.
Tel. +33.6.33 17 20 85.
E-mail: mprodel@hevaweb.com

References

1. Morlat P. [Prise en charge du VIH - Recommandations du groupe d'experts. Conseil national du sida et des hépatites virales].[in French]. 2019. Accessed: 15 april 2021. Available from: https://cns.sante.fr/actualites/prise-en-charge-du-vih-recommandations-du-groupe-dexperts

2. Silue Y. [Surveillance des infections à VIH et Sida en Ile-de-France].[in French]. Bull Veille Sanit; 2015.

3. de Léotoing L, Yazdanpanah Y, Finkielsztejn L, et al. Costs

associated with hospitalization in HIV-positive patients in France. AIDS 2018;32:2059−66.

4. Trapero-Bertran M, Oliva-Moreno J. Economic impact of HIV/AIDS: a systematic review in five European countries. Health Econ Rev 2014;4:15.

5. Krentz HB, Vu Q, Gill MJ. Updated direct costs of medical care for HIV-infected patients within a regional population from 2006 to 2017. HIV Med 2020;21;289−98.

6. Ong KJ van Hoek AJ, Harris RJ, et al. HIV care cost in England: a cross-sectional analysis of antiretroviral treatment and the impact of generic introduction. HIV Med 2019;20:377−91.

7. Guillon M, Celse M, Geoffard PY. Economic and public health consequences of delayed access to medical care for migrants living with HIV in France. Eur J Health Econ 2018;19:327−40.

8. Perelman J, Rosado R, Amri O, et al. Economic evaluation of HIV testing for men who have sex with men in community-based organizations – results from six European cities. AIDS Care 2017;29:985−9.

9. Durand-Zaleski I, Mutuon P, Charreau I, et al. Costs and benefits of on-demand HIV preexposure prophylaxis in MSM. AIDS 2018;32:95−102.

10. Demessine L, Peyro-Saint-Paul L, Gardner EM, et al. Risk and cost associated with drug–drug interactions among aging HIV patients receiving combined antiretroviral therapy in France. Open Forum Infect Dis 2019;6;ofz051.

11. Cohen JP, Beaubrun A, Ding Y, et al. Estimation of the incremental cumulative cost of HIV compared with a non-HIV population. Pharmacoecon Open 2020;4:687−96.

12. Mostardt S, Hanhoff N, Wasem J, et al. Cost of HIV and determinants of health care costs in HIV-positive patients in Germany: results of the DAGNÄ K3A study. Eur J Health Econ 2013;14:799−808.

13. Bezin J, Duong M, Lassalle R, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2017;26:954−62.

14. Cavailles A, Melloni B, Motola S, et al. Identification of patient profiles with high risk of hospital re-admissions for acute COPD exacerbations (AECOPD) in France using a machine learning model. Int J Chron Obstruct Pulmon Dis 2020;15:949−62.

15. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead.

Nat Mach Intell 2019;1:2016-15.

16. Shivade C, Raghavan P, Fosler-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc 2014;21:221−30.

17. Solares JRA, Raimondi FED, Xhu Y, et al. Deep learning for electronic health records: A comparative review of multiple deep neural architectures. J Biomed Inform 2020;101:103337.

18. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med 2019;25:24-9.

19. Oliveira HD, Prodel M, Augusto V. Binary classification on French hospital data: Benchmark of 7 machine learning algorithms. 2018 IEEE Int Conf on Systems, Man, and Cybernetics (SMC), 2018, p. 1743−1748.

20. Ghassemi M, Naumann T, Schulam P, et al. A review of challenges and opportunities in machine learning for health. AMIA Jt Summits Transl Sci Proc 2020;2020:191-200.

21. Breiman L, Friedman J, Olshen R, Stone CJ. Classification and regression trees. Biometrics 1984;40:874.

22. Wong J, Horwitz MM, Zhou L, Toh S. Using machine learning to identify health outcomes from electronic health record data. Curr Epidemiol Rep 2018;5:331−42.

23. Caisse nationale d'Assurance maladie des Travailleurs salariés. [Améliorer la qualité du système de santé et maîtriser les dépenses].[in French]. Paris Caisse Nationale Assurance Maladie, 2018. Available from: https://www.vie-publique.fr/rapport/29982-ameliorer-la-qualite-du-systeme-de-sante-et-maitriser-les-depenses

24. Katlama C, Ghosn J, Murphy RL. Individualized antiretroviral therapeutic approaches: less can be more. AIDS 2017;31: 1065−71.

25. Sloan CE, Champenois K, Choisy P, et al. Newer drugs and earlier treatment: Impact on lifetime cost of care for HIV-infected adults. AIDS 2012;26:45-56.

26. Morlat P, Roussillon C, Henard S, et al. Causes of death among HIV-infected patients in France in 2010 (national survey): trends since 2000. AIDS 2014;28:1181−91.

27. May T, Lewden C, BonnetF, et al. [Causes and characteristics of death among HIV-1 infected patients with immunovirologic response to antiretroviral treatment].[Article in French]. Presse Med 2004;33:1487-92.

28. Meffre C, Le Strat Y, Delarocque-Astagneau E, et al. Prevalence of hepatitis B and hepatitis C virus infections in France in 2004: Social factors are important predictors after adjusting for known risk factors. J Med Virol 2010;82:546-55.