Predictive ranges for the population of CD4 + lymphocytes for HIV positive patients on antiretroviral treatment

Javier Rodríguez[1], Carlos Pérez[2], and Ribká Soracipa[1]

[1]Insight Research Group SAS
[2]Servicios y Asesorías en Infectología

March 30, 2022

Abstract

Introduction: part of the effectiveness of follow-ups of patients with HIV in antiretroviral therapy is done through the quantification of CD4 + lymphocytes, hence the correct establishment of these values is an issue of interest in the clinical setting. Objective: to establish predictive mathematical relationships between CD4+ cell counts in ranges >500, [200,500], <200, between 200 and >500 and <200 up to 500 cells/μL3 with the absolute leukocyte count of patient samples over time in the context of the theory of probability. Methods: Through an inductive process carried out in 11 patient samples, mathematical patterns that forecast in time the correspondence between absolute leukocyte counts and CD4+ counts that can occur in five ranges of clinical interest. Then, a confirmation was done with 139 patients in a blind study obtaining the probability values for each range as well as sensitivity and specificity. Results: The five dynamics predicted achieved probabilities that varied between 0.96 and 1, with a global probability of 0.99 with sensitivity and specificity values of 99%. Conclusions: a self-organized mathematical temporal order that allows to forecast the values of CD4+ cells in relation to leukocyte counts in ranges of clinical interests was found, which could be useful to develop surveillance programs of HIV-infected patients in low-income countries, improving their survival rates.

Introduction

The World Health Organization (WHO) recommends that antiretroviral therapies initiate according to the $CD4^+$ cells count, among other clinical considerations (1), as has been established by decades (2). The timely beginning of antiretroviral therapy contributes to suppress viral replication and augmenting immune response most effectively. Therefore, part of the following up of patients is done through the measurement of this variable (3,4).

Some of the investigations conducted in this field, have analyzed the behavior of $CD4^+$ counts when antiretroviral treatment has initiated (2,5) through descriptive and randomized clinical studies (6,7). Among the strategies that have been included in these studies is the implementation of models with lineal-mixed effects (7-9), which exhibit a lineal tendency in the trajectory of $CD4^+$cell counts, however, there is a lack of correlation between these dynamics and the real ones of the patients. Considering these limitations, the application of asymptotic non-lineal mixed effects models of three parameters, that can be viral load, age and sex, have allowed the analysis of non-lineal trajectories of $CD4^+$ cells (10,11).

In the history of mathematics, probability achieved the level of theory when axioms through operations between sets were established (12). Among the rules of sets, the concept of belonging is found, which establishes that a set is defined by the elements that belong to that set, which is similar in nature to a sample or probability space that groups the possible events of a determined experiment (12). From these mathematical bases, methodologies that forecast the behavior of $CD4^+$ cells in specific ranges were developed from leukocytes and lymphocytes counts, relying only on the complete blood count (13,14). The unification

of these theories refined previously defined ranges to characterize the behavior of $CD4^+$cells when these are inferior to normal counts (13,14), which could optimize the beginning of antiretroviral therapy.

The purpose of this investigation is to establish predictive mathematical relationships in time for the counts of $CD4^+$ cells through leukocyte counts analyzing $CD4^+$ counts in ranges >500 cells/μL^3, between 200 to 500 cells/μL^3, <200 cells/$\mu L^3$and the oscillations between 200 and >500 cells/μL^3 and <200 up to 500 cells/μL^3 in the context of the theory of probability.

Methodology

Data source

Different registries in time of 150 patients were analyzed, observing the behavior of absolute counts of leukocytes, metamyelocytes, neutrophils, lymphocytes, monocytes, eosinophils, basophils and reactive lymphocytes; percentage of lymphocytes, leukocytes, metamyelocytes, neutrophils, lymphocytes, monocytes, eosinophils, basophils and reactive lymphocytes; erythrocyte sedimentation rate, mean platelet value and red blood count.; serology (RPR), RNA, HIV-1, viral load ($\log_{10}$); $TCD3^+$, $TCD4^+$ and $TCD8^+$ lymphocytes; $CD4^+/CD8^+$ ratio and total lymphocytes; scattergram; platelet distribution width; mean corpuscular volume; prolymphocytes; mean corpuscular hemoglobin; mean corpuscular hemoglobin concentration; red blood cell distribution width; blast cells; promyelocytes; myelocytes. Variables such as sex, age, years of diagnosis, antiretroviral therapy, among others, were not considered.

The values of these variables were taken from a database analyzed by a specialist in infectious diseases that was previously developed by the enterprise "Servicios y Asesorías en infectología"

Procedure

A mathematical induction with 11 samples, denominated prototypes, was performed in order to obtain mathematical relationships between both total leukocytes and $CD4^+$ cell counts. The study of these relationships yielded mathematical patterns, defining 5 possible dynamics as follows:

1. Dynamics in which all the registries of the sequence's present values of $CD4^+$ lymphocytes >500 cells/μL^3.
2. Dynamics in which all the registries of the sequence's present values of $CD4^+$ lymphocytes between [200,500] cells/μL^3.
3. Dynamics in which all the registries of the sequence's present values of $CD4^+$ lymphocytes lesser than 200 cells/μL^3.
4. Dynamics in which the registries of the sequence's present values of $CD4^+$ lymphocytes >500 cells/μL3 but also between [200,500] cells/μL^3.
5. Dynamics in which the registries of the sequence's present values of $CD4^+$ lymphocytes <200 cells/μL3 but also between [200,500] cells/μL^3.

Given these 5 dynamics, predictive mathematical parameters were established and were applied to the remaining 139 cases through a blind study in which a software developed by Insight Group in C^{++} was used to conduct all the predictive calculations. It is worth noting that the methodology follows the steps of probability in quantum mechanics and it is not the blind study that allows to obtain predictions in time.

Ethical aspects

This study follows the scientific, technical, and administrative regulations for investigation in health of the Colombian Health Ministerium through resolution 8430 of 1993, specially in title 11 referring to investigation in human beings, considering it as a no risk investigation given that mathematical calculations are performed over previously recollected results of clinical test that had been prescribed to patients according to clinical protocols. The integrity and anonymity of patients was respected during the investigation since personal data was not included in the database and no diagnostic or therapeutic interventions were performed with them. This investigation is also based on the ethical principles for investigation that involucrate human beings of the World Medical Association's Declaration of Helsinki, the Nuremberg Code, and the Belmont Report.

Results

The mathematical induction performed in the 11 prototypes gave the following results, for a CD4+/μL^3 count that is in the range of [CD4+ >500] cells/μL^3, the total values varied between 503 y 858 and it has a white blood cell count/μL^3 that varied between 3.6 y 6.5; the count CD4+/μL^3 that is in the range of [200< CD4+<500] cells/μL^3 varied between 249 y 489, with a leucocyte count/μL^3 that varied between 3.3 y 5.8, the CD4+/μL^3 count is in the range [CD4+ <200] cells/μL^3 varied between 80 y 147, has a range leucocytes/μL^3 that varied between 2.9 y 6.3 (see table 2). The difference in days from the first sample taken for the prototype patients and the next one varied between 133 y 950 days (see table 2).

609 registries systematized in different dates of 150 patients were analyzed, observing the values of $CD4^+$ cells and total leukocyte counts. In table 1, the information of the most representative patients regarding the quantity of registries in ranges of $CD4^+$ cells >500 cells/μL^3, between 200 and 500 cells/μL^3 and <200 cells/μL^3 are exhibited as well as the results of the predictions. For the first range, 22 patients had 1 registry, 18 patients had 2 registries, 22 patients had 3 registries, 26 patients had 4 registries, 17 patients had 5 registries and 6 patients had 6 registries, while for the intermediate range 24 had 1 registry, 18 patients had 2 registries, 12 patients had 3 registries, 22 patients had 4 registries, 9 patients had 5 registries and 1 patient had 6 registries and for the last range, 9 patients had 1 registry, 3 patients had 2 registries, 2 patients had 6 registries and none had 3, 4 and 6 registries (table 1). It is worth noting that the repetition of registries and the values among those patients that had fluctuations in the ranges between 200 and >500 cells/μL^3as well as <200 up to 500 cells/μL^3 was also considered.

To observe the behavior in time of $CD4^+$ and total leukocytes counts, scatterplots were designed, with the variable time in X axis and the values of the cells in Y axis (see figures 1 and 2).

Predictive results

The mathematical induction conducted with the 11 prototypical cases yielded 8 different mathematical patterns according to the defined ranges and their combinations:

1. Two registries are taken, and it is looked whether one of the following configurations is presented:
2. One of the measurements presents $CD4^+$ lymphocytes >500 cells/μL^3 and the other between [200, 500] cells/μL^3.
3. One of the measurements presents $CD4^+$ lymphocytes between [200, 500] cells/μL^3 and the other <200 cells/μL^3.
4. Both have $CD4^+$ lymphocytes >500 cells/μL^3 and leukocytes >3.7 cells/mm^3.
5. Both have a population of $CD4^+$ lymphocytes <200 cells/μL^3 and measurements of leukocytes between [2.0, 3.9] cells/mm^3.
6. Both have a population of $CD4^+$ lymphocytes between [200, 500] cells/μL^3 and at least one of the leukocyte measurements presents values [?] 4 cells/mm^3.
7. If case *a* is presented, and the values of leukocytes are between [3.0, 3.9] cells/mm^3 and the following measurement is within that range, then the associated $CD4^+$ populations will be >500 cells/μL^3 or will be [200, 500] cells/μL^3.
8. If case *a* is presented and the measurement that presents a value of $CD4^+$ between [200, 500] cells/μL^3 also present a value of leukocytes [?] 4 cells/mm^3 and the measurement with a value of >500 cells/μL^3 also presents a value of leukocytes [?] 3.7 cells/mm^3, then the most likely event is that $CD4^+$ values are between [200, 500] or >500 cells/μL^3.
9. If case *b* is presented, and the value of leukocytes in the measurement that presents $CD4^+$ values <200 cells/μL^3 is <3 cells/mm^3, and for the measurement that presented a value of $CD4^+$ lymphocytes between [200, 500] cells/μL^3 has a leukocyte value [?] 4 cells/mm^3, if in the following measurement a value of leukocytes is between [4,6] cells/mm^3 the most likely event is that the value of $CD4^+$ is found between [200, 500] cells/μL^3 but if the value of leukocytes is >6 a value of $CD4^+$<200 cells/μL^3 can be found.
10. If case *b* is presented, and the value of leukocytes in the measure that contains $CD4^+$ <200 cells/μL^3

is higher than 3 cells/mm^3, and for the registry of CD4$^+$ between [200,500] cells/μL^3 a measure of leukocytes lesser than 3 cells/mm^3 is presented, it is more likely that if the value of leukocytes is higher than 3 cells/mm^3, then the measurement of CD4$^+$ will be between [200,500] cells/μL^3.

11. If case c is presented, the greater probability is that in the posterior measurements, when leukocyte populations are [?] 3,7 cells/mm^3, the associated CD4$^+$populations will be >500 cells/μL^3.
12. If case *d* is presented, the most likely event is that when a value of leukocytes between 2 and 3 cells/mm^3 in the following measurement, the associated CD4$^+$populations will be <200 cells/μL^3.
13. If case *e* is presented, the most likely event is that if in the following measurements leukocytes values are [?] 4 cells/mm^3, the associated CD4$^+$populations will be [200, 500] cells/μL^3.

The overall predictive precision of the mathematical predictive patterns was 0.99 (see table 3) and the values of sensitivity and specificity obtained in the blind study were 99%.

Table 1. Quantities of registries by ranges of the most representative cases analyzed for the study and result of the prediction for each case.

Case	Χυαντιτψ οφ ρεγιστριες βψ ρανγες (ςΔ4$^+$/μΛ^3)	Χυαντιτψ οφ ρεγιστριες βψ ρανγες (ςΔ4$^+$/
	>500	**200-500**
1	4	
2	4	
3	5	
4	4	
5	2	
6	6	
7	3	
8	2	
9	5	
10	4	
63		3
64		4
65		4
66		2
67		4
68		5
69		4
70		4
71		4
72		4
88		
89		

Where S: fulfills the diagnostic predictive criteria; N: does not full the diagnostic predictive criteria.

Table 2. Values of CD4$^+$ cell counts, Dif: Difference in days between the first sample and the next sample taken; CD4$^+$ ranges, leukocytes counts and dates of the prototypes selected for mathematical induction.

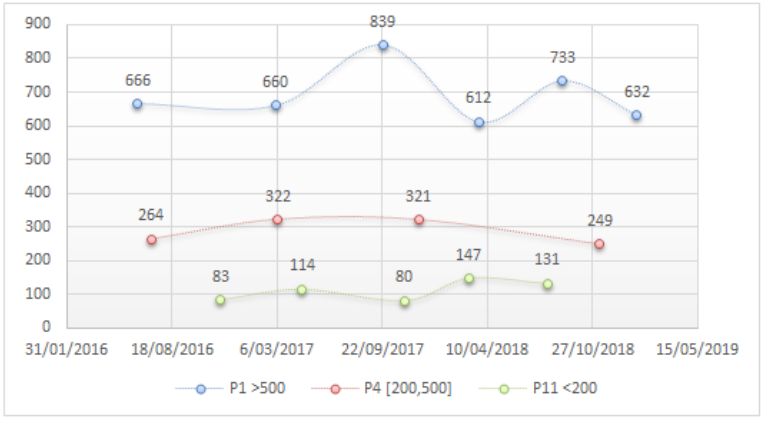
Protype	Date	Dif	ςΔ4$^+$/μΛ^3	CD4+ range	Λευκοςψτες$^+$/μΛ^3
P1	13/06/16		666	>500	5.6
	03/03/17	263	660	>500	5
	23/09/17	467	839	>500	5.2
	24/03/18	649	612	>500	5.9

Protype	**Date**	**Dif**	**ˆΔ4^{+}/μΛ^{3}**	**CD4+ range**	**Λευκοςψτες^{+}/μΛ^{3}**
	31/08/18	809	733	>500	4.5
	19/01/19	950	632	>500	4.6
P2	07/10/16		687	>500	6.3
	30/06/17	266	858	>500	5.5
	14/02/18	495	590	>500	5.7
	06/08/18	668	687	>500	4.3
P3	05/10/16		805	>500	5.4
	24/04/17	201	745	>500	6
	18/10/17	378	702	>500	5.3
	05/04/18	547	692	>500	4.8
P4	09/07/16		264	[200-500]	5.8
	07/03/17	241	322	[200-500]	5.2
	30/11/17	509	321	[200-500]	5
	09/11/18	853	249	[200-500]	5.5
P5	01/10/16		278	[200-500]	3.4
	14/02/17	136	362	[200-500]	3.3
	31/07/17	303	401	[200-500]	4.6
	23/02/18	510	424	[200-500]	3.7
P6	05/10/16		758	>500	4.4
	26/04/17	203	503	>500	4.5
	18/11/17	409	625	>500	6.5
	26/11/18	782	411	[200-500]	4.3
P7	27/07/16		489	[200-500]	5.3
	07/01/17	164	530	>500	5.9
	18/07/17	356	528	>500	5.2
	30/01/18	552	576	>500	6.1
	10/01/19	897	477	[200-500]	5.3
P8	30/07/16		482	[200-500]	3.5
	10/12/16	133	428	[200-500]	3.7
	12/06/17	317	368	[200-500]	3.8
	09/01/18	528	642	>500	3.6
	07/12/18	860	453	[200-500]	3.3
P9	12/11/16		464	[200-500]	4.6
	29/06/17	229	601	>500	4.6
	07/02/18	452	400	[200-500]	4.8
	08/02/19	818	682	>500	5.7
P10	10/09/16		645	>500	5.5
	22/04/17	224	743	>500	6.1
	23/10/17	408	547	>500	5.7
	21/12/18	832	470	[200-500]	4.9
P11	18/11/2016		83	<200	4.4
	22/04/2017	155	114	<200	4.5
	3/11/2017	350	80	<200	4.1
	5/03/2018	472	147	<200	6.3
	2/08/2018	622	131	<200	4

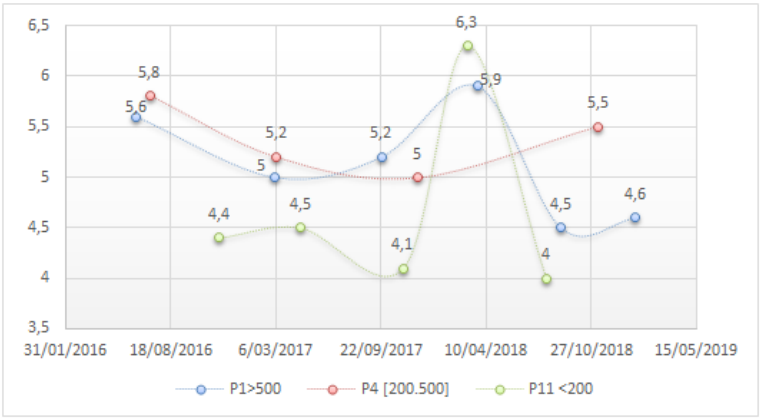
Tabla 3. Values of probability (P) obtained for all the ranges evaluated with the methodology

Ranges evaluated	P
[CD4+ > 500]	1
[200< CD4+ < 500]	0,96
[CD4+ < 200]	1
[CD4+ > 500] y [200<CD4+ < 500]	0,98
[200 < CD4+ < 500] y [CD4+ < 200]	1
Totality	0,99

Figure 1 . Behavior in time of a) $CD4^+$ cell counts for the ranges <200, between 200 to 500 and >500 cells/μL^3 and b) leukocyte counts in cells/μL^3 for the cases of patient No. P1, P 4 and P11 of the prototypical cases.

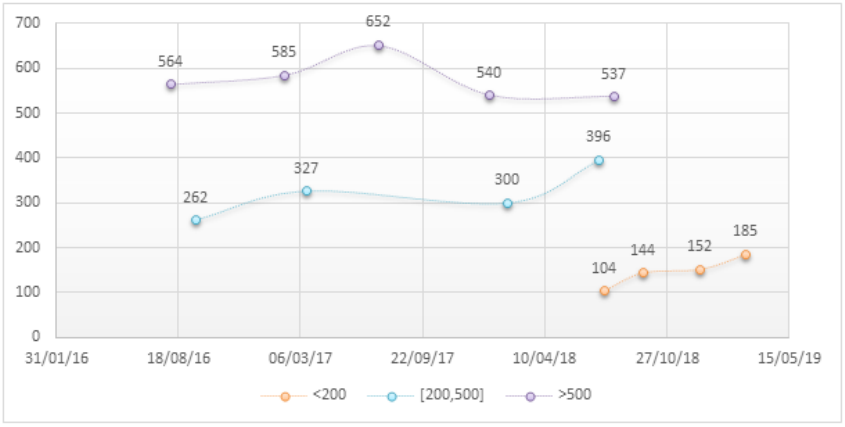


a)

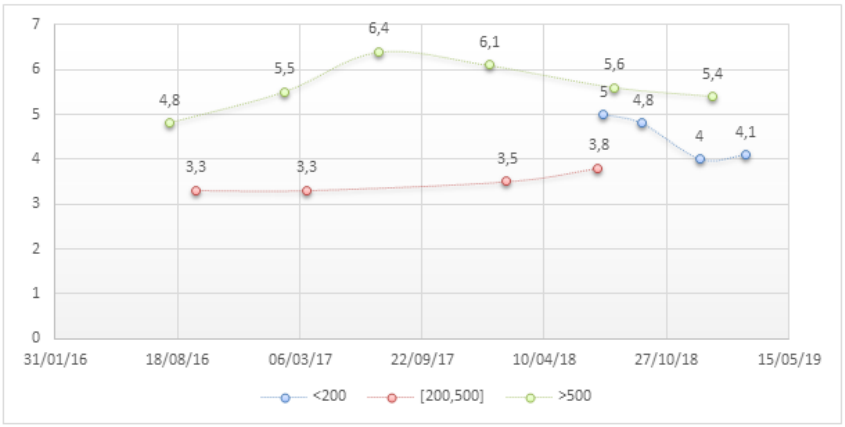


b)

Figure 2 . Representative dynamics of $CD4^+$lymphocytes (a) and leukocytes (b) of the cases 36, 90 and 117 with counts <200, between the range [200, 500] and >500.



a)



b)

Discussion

This is the first investigation that established predictive mathematical relationships in five dynamics of the $CD4^+$ counts in HIV-infected patients in antiretroviral treatment through probability theory and total leukocyte counts. The results of the mathematical induction are proof of the capacity of theoretical physics to formulate theories that contribute and the understanding of phenomena in nature such as the relationship between $CD4^+$ cells and leukocytes counts. The results of probability obtained, that varied between 0.96 and 1, with an overall probability of 0.99, strongly suggest that this phenomenon is deterministic as well as the sensitivity and specificity values obtained of 99% indicate its clinical usefulness.

Likewise, these results also suggest that the methodology could be helpful to conduct surveillance of patients in low-income countries, where flow cytometry is not readily available given its elevated costs and requirement of specialized equipment and personnel, since physicians would only require to observe the patterns of leukocytes in complete blood counts to obtain the most likely value of the range of $CD4^+$ cells in time for each patient which in turn allows to change antiretroviral regimens or addition antibiotics for the prevention of opportunistic infections, according to the values obtained, improving the overall mortality in these patients.

Measurements of CD4+/μL^3 and viral load, for example, require equipment with high technology, which determines the cost that this laboratory test can have (15). Families that are in geographical areas such as central Africa, would not have the economic resources that allow them to cover the cost of these tests, and even more so when it comes to laboratory tests that need to be practiced on a regular basis (15,16). Routine

tests such as complete blood count are the best option for these families (15), so the results of this study are perfectly applicable to the results of this type of examination.

It is worth noting that in the literature there are clinical studies, which have sought to establish another type of relationship between variables to establish CD4 counts (15,17). Variations in the results of these studies are associated with the patient's race and sex (21), environmental influences and sample size are also considered as causes of fluctuations in the values of CD4 counts (15). In a study conducted to determine mortality predictions from the relationship between CD4 count and viral load, he found that predictions of mortality are more reliable from the CD4 count, and the variation in this count is due to the changes presented in the viral load and the duration of antiretroviral treatment (18). However, it is a study based on a linear model, which was broken at the time of correlating the results with the real data, which implies more studies until the parameters established in this study are correlations with the real data. On the other hand, the present study is not based on a model that seeks to simulate the dynamics of the CD4 + count, but to establish from a theoretical basis general values applicable to all types of patients.

This study also provides an advantage over the previous considerations, since it does not rely on populational factors, which usually complicate the analysis, especially when dynamics want to be predicted, contributing in an objective and reproducible manner (2,5) to the approximation of this problem, which has not been dilucidated in other clinical conventional studies (6,7).

Other methodologies with predictive capacity designed in the line of theoretical physics, can be observed in the studies conducted in cardiology through mathematical and geometrical parameters that differentiate normal from abnormal dynamics (19), also achieving predictions of mortality in the intensive care unit (20), the number of people infected by malaria (21) and the people living with HIV/AIDS (22).

Acknowledgments

We extend our thanks to Doctor Marco Fonseca, president of the Asociación Colombiana de Neurocirugía and Doctor Germán Forero, Investigations Director, for their support to our investigations.

References

1. Kebede M, Zegeye DT, Zeleke BM. Predicting CD4 count changes among patients on antiretroviral treatment: Application of data mining techniques. Comput Methods Programs Biomed. 2017; 152:149-157.
2. Means AR, Risher KA, Ujeneza EL, Maposa I, Nondi J, Bellan SE. Impact of Age and Sex on $CD4^+$ Cell Count Trajectories following Treatment Initiation: An Analysis of the Tanzanian HIV Treatment Database [published correction appears in PLoS One. 2017 Jan 6;12 (1):e0170155]. *PLoS One* . 2016;11(10):e0164148. Published 2016 Oct 7. doi:10.1371/journal.pone.0164148.
3. Ratnam M, Nayyar AS, Reddy DS, Ruparani B, Chalapathi KV, Azmi SM. CD4 cell counts and oral manifestations in HIV infected and AIDS patients. J Oral Maxillofac Pathol. 2018;22(2):282.
4. Gaurav S, Keerthilatha PM, Archna N. Prevalence of oral manifestations and their association with CD4/CD8 ratio and HIV viral load in South India. Int J Dent 2011. 2011 964278.
5. Lok JJ, Bosch RJ, Benson CA, et al. Long-term increase in $CD4^+$ T-cell counts during combination antiretroviral therapy for HIV-1 infection. AIDS. 2010;24(12):1867–1876. doi:10.1097/QAD.0b013e32833adbcf
6. Luz PM, Grinsztejn B, Velasque L, Pacheco AG, Veloso VG, Moore RD, et al. Long-Term $CD4^+$ Cell Count in Response to Combination Antiretroviral Therapy. PloS one. 2014;9(4).
7. Wright ST, Petoumenos K, Boyd M, Carr A, Downing S, O'Connor CC, et al. Ageing and long-term CD4 cell count trends in HIV-positive patients with 5 years or more combination antiretroviral therapy experience. HIV Medicine. 2013;14(4):208–16.
8. Hoffmann CJ, Schomaker M, Fox MP, et al. CD4 count slope and mortality in HIV-infected patients on antiretroviral therapy: multicohort analysis from South Africa. J Acquir Immune Defic Syndr. 2013;63(1):34–41.

9. Kye-Hyung Kim, Jongyoun Yi, Sun Hee Lee. The CD4 slope can be a predictor of immunologic recovery in advanced HIV patients: a case-control study. Korean J Intern Med 2015;30(5):705-713.
10. De Beaudrap P, Etard JF, Diouf A, Ndiaye I, Gueye NF, Gueye PM, et al. Modeling $CD4^+$ cell count increase over a six-year period in HIV-1-infected patients on highly active antiretroviral therapy in Senegal. The American journal of tropical medicine and hygiene. 2009;80(6):1047–53.
11. Lewis J, Walker AS, Castro H, De Rossi A, Gibb DM, Giaquinto C, et al. Age and CD4 count at initiation of antiretroviral therapy in HIV-infected children: effects on long-term T-cell reconstitution. The Journal of infectious diseases. 2011;205(4):548–56.
12. Gnedenko B, Khintchine, A. Introducción a la teoría de las probabilidades. Barcelona: Montaner y Simon. 1968.
13. Rodríguez J, Prieto S, Correa C, Mora J, Bravo J, Soracipa Y, et al. Predictions of CD4 lymphocytes' count in HIV patients from complete blood count. BMC Medical Physics. BMC Medical Physics. 2013; 13:3.
14. Rodriguez J, Prieto S, Correa C, Melo M, Dominguez D, Olarte N, Suárez D, Aragón L, Torres F, Santacruz F. Prediction refinement of CD4 cells number based on sets and probability theory. Current HIV Research. 2018; 16: 416-424.
15. Mbanya D, Assah F, Ndembi N, Kaptue L. Monitoring antiretroviral therapy in HIV/AIDS patients in resource-limited settings: CD4 counts or total lymphocyte counts? Int J Infect Dis. 2007 Mar;11(2):157-60.
16. Ouattara EN, Robine M, Eholié SP, et al. Laboratory Monitoring of Antiretroviral Therapy for HIV Infection: Cost-Effectiveness and Budget Impact of Current and Novel Strategies. Clin Infect Dis. 2016;62(11):1454–1462. doi:10.1093/cid/ciw117
17. Brennan AT, Maskew M, Sanne I, Fox MP. The interplay between CD4 cell count, viral load suppression and duration of antiretroviral therapy on mortality in a resource-limited setting. Trop Med Int Health. 2013;18(5):619–631. doi:10.1111/tmi.12079.
18. Gorter RW, Vranizan KM, Osmond DH, Moss AR. Differences in laboratory values in HIV infection by sex, race, and risk group. AIDS. 1992 Nov;6(11):1341-7..
19. Rodríguez J, Prieto S, Ramírez L. A novel heart rate atractor for the prediction of cardiovascular disease. Informatics in medicine. 2019; 15(100174):1-9.
20. Rodríguez J. Dynamical systems applied to dynamic variables of patients from the Intensive Care Unit (ICU). Physical and mathematical Mortality predictions on ICU. J. Med. Med. Sci. 2015; 6(8): 102-108.
21. Rodríguez J. Spatio-temporal probabilistic prediction of appearance and duration of malaria outbreak in municipalities of Colombia. J. Phys.: Conf. Ser. 2019; 1160:1-7.
22. Rodríguez J, Oliveros D, Soracipa Y, Bernal L, Correa C, Abrahem L, Flórez L, Rodríguez D, Bahamon J, Valero O. Análisis probabilista con caminata al azar del número de personas viviendo con VIH mundialmente. Rev. Fac. Nac. Salud Pública. 2018; 36(1):27-33.