ORIGINAL PAPER

Mathematical predictive relationship of CD4+ lymphocytes and total leukocytes in HIV-infected patients

Signed Esperanza Prieto Bohorquez[1], Joao Cuesta[2], César Alejandro Valdés Cadena[1], Carlos Eduardo Pérez Díaz[3,4], Sandra Medina Araujo[5], Ribká Soracipa Muñoz[1], Andrés David Ibarra[6], Eduardo Méndez Pino[6], Sandra Catalina Correa Herrera[1], Javier Oswaldo Rodríguez Velasquez[6]

[1]Insight Group, Nacional University Hospital of Colombia, Bogotá, Colombia
[2]Humanitas Group, Faculty of Education and Humanities, Universidad Militar Nueva Granada, Bogotá, Colombia
[3]Services and Consulting in Infectology, Bogotá, Colombia
[4]Marly Clinic, Bogotá, Colombia
[5]Faculty of Basic and Applied Sciences, Universidad Militar Nueva Granada, Bogotá, Colombia
[6]Faculty of Medicine, Universidad Militar Nueva Granada, Bogotá, Colombia

Abstract

Introduction: Different parameters have been established to direct the treatment of patients with human immunodeficiency virus (HIV) infection, such as CD4+ lymphocyte values, and it is of clinical interest to have methodologies that accurately predict these values. Aim of the study was to predict the total values of leukocytes and CD4+ lymphocytes greater than 500 cells/μl^3 in HIV-infected patients using the theory of probability and set theory.

Material and methods: Starting from 7 cases with several records over time, an induction was performed establishing mathematical patterns between CD4+ lymphocyte values and total leukocyte values, while applying the probability theory to calculate predictive accuracy in 43 cases, and subsequently, sensitivity and specificity were calculated in a blinded study.

Results: In total, 184 records were analyzed for 50 cases. The values of total leukocytes equal to or greater than 3.9 cells/mm^3 were predicted to correspond to CD4+ lymphocyte values greater than 500 cells/μl^3 in 100% of time, with sensitivity and specificity results of 100%.

Conclusions: This is the first investigation with the theory of probability, in which predictions were made from leukocyte values equal to or greater than 3.9 cells/mm^3 to find CD4+ lymphocyte counts. A predictive probabilistic methodology was developed, and determined results for the calculated ranges were found.

HIV AIDS Rev 2024; 23, 2: 124-129
DOI: https://doi.org/10.5114/hivar/149277

Key words: probability, leukocyte count, HIV, CD4+ lymphocyte count.

Address for correspondence: Dr. Signed Esperanza Prieto Bohórquez, Insight Group. Hospital Universitario Nacional de Colombia, Bogotá, Colombia, e-mail: insightgroup.ext@gmail.com

Article history:
Received: 26.09.2021
Received in revised form: 06.03.2022
Accepted: 19.04.2022
Available online: 21.05.2024

International Journal of HIV-Related Problems
HIV & AIDS Review

Introduction

Probability can be understood as a dimensionless measure of obtaining the possibility of occurrence of an event in the future, for which axiomatic conditions defined by Kolmogorov test must be fulfilled and all probabilities are considered to have a value between 0 and 1, where 1 is an event that will surely occur, with the entire space showing a probability of 1. In case of two events considered, the probability will result from the sum of both [1-4].

According to the Joint United Nations Program (UNAIDS), there are about 36 million people currently living with HIV infection [5]. This virus has a tropism for cell lines, such as macrophages or CD4+ T lymphocytes [6], and as the values of these cells decrease, the risk of developing opportunistic infectious diseases increases [7].

Different methods of follow-up have been developed for patients with human immunodeficiency virus (HIV) infection, including the measurement of CD+ 4 T lymphocyte level; however, the cost associated with this measurement using flow cytometry limits the possibility of timely follow-ups, during which prophylactic treatments can be modified, particularly in Third World countries [8-10]. Considering this difficulty, alternatives to obtain predictions of CD4+ T lymphocyte counts through different methodologies, such as machine learning, artificial neural networks, and linear analysis, are used. However, currently, there is no high precision results to allow an exploration of this phenomenon [11-14].

Similarly, from the theory of probability and theory of sets, a methodology has been designed that relates the variables of hemogram, such as the absolute counts of leukocytes and lymphocytes to predict the absolute CD4+ counts. For this, ranks of 1,000 cells/ml^3 analyzed by intersections were organized, obtaining percentages of 100% predictions for counts of less than 4,000 leukocytes [15, 16]. This proves that the methods of exploration of problems from mathematics can generate clinical solutions [15, 16].

The objective of this study was to predict the temporal dynamics of absolute CD4+ counts from the absolute leukocyte counts of blood count by establishing mathematical relationships that describe this phenomenon initially with a mathematical induction, and later with a blinded study of 50 cases.

Material and methods

CD4 + range > 500 cells/µl^3 bound to total leukocyte counts, in which seven prototype cases (P1 to P7) were taken to generate a mathematical induction, considering minimums and maximums of the counts.

Population

There were 2 to 6 flow cytometry records and blood count of 50 patients with confirmed HIV infection, with CD4+ lymphocyte counts greater than 500 cells/µl^3 at different dates regardless of age, sex, or treatment from baseline data, evaluated by an infectious disease specialist from Servicios y Asesorías en Infectogía, and different records were entered between 2016 to 2019.

Procedure

Initially, a mathematical induction was performed for 7 of the most representative cases, of which the absolute counts of leukocytes and CD4+ lymphocytes were observed in search of a mathematical order between these two variables, with predictions that could be made from CD4+ counts.

First, this inductive process consisted of simultaneous observation of experimental behavior of CD4+ T lymphocyte counts with respect to the total leukocyte counts with a significant behavior. This means that cases that presented extreme values of both variables were taken to generate 7 prototypes, which included possible scenarios of the studied dynamic. Subsequently, an analogy based on quantum electrodynamics was made, where the probability of the photon was established based on probabilities of the positron and electron when converging. In this case, the analogy was directed to establish the probability of CD4+ counts from the convergence of probabilities of both CD4+ and total leukocytes counts. Once the observation of experimental data and physical analogy was made, a predictive dynamic was established, from which CD4+ counts greater than 500 cells/µl^3 were predicted based solely on total leukocyte count value. This was a value of interest considered in clinical context.

It is important to note that in theoretical physics, inductive predictive processes allow predicting universal phenomena based on scarce experimental observations. In this case, the 7 cases analyzed served as the basis for inductive process, in which predictive parameter was automated by means of a software developed by Insight Group in C++ language, to verify the precision of the parameter in a blinded study with the remaining cases not analyzed.

Therefore, the total leukocyte values of the remaining 43 cases were taken to develop a blinded study, and predictions of CD4+ counts greater than 500 cells/µl^3 were generated. For this, values of leukocytes were considered, and by means of the predictive parameter generated in mathematical induction and automated in software, values of CD4+ counts were predicted. Finally, methodological predictive precision was verified using calculations of the probability of presentation of the range.

Statistical analysis

Values of CD4+ lymphocytes were unblinded to perform sensitivity and specificity calculations using a binary table that included true positives and negatives as well as false positives and negatives.

Ethical aspects

This research followed scientific, technical, and administrative guidelines outlined for health research in resolution

Table 1. Total CD4+ counts of 50 cases analyzed and time intervals between records in number of days

Case	No. of records	Record 1	Record 2	Record 3	Record 4	Record 5	Record 6
1	1	25/02/19	–	–	–	–	–
2	1	23/10/17	–	–	–	–	–
3	2	10/10/17	363	–	–	–	–
4	2	12/10/18	159	–	–	–	–
5	2	25/11/16	224	–	–	–	–
6	2	07/03/18	275	–	–	–	–
7	2	01/03/18	217	–	–	–	–
8	2	06/04/18	208	–	–	–	–
9	2	22/11/16	527	–	–	–	–
10	3	13/12/16	176	199	–	–	–
11	3	02/12/16	136	541	–	–	–
12	3	13/05/17	171	329	–	–	–
13	3	24/04/17	143	453	–	–	–
14	3	09/11/16	224	162	–	–	–
15	3	25/10/16	637	136	–	–	–
16	3	30/11/17	167	197	–	–	–
17	3	02/02/18	253	153	–	–	–
18	4	06/12/16	172	178	155	–	–
19	4	19/07/17	171	186	182	–	–
20	4	15/11/16	164	208	204	–	–
21	4	16/11/16	204	198	381	–	–
22	4	08/11/16	192	313	246	–	–
23	4	16/01/17	170	149	337	–	–
24	4	25/05/17	139	182	182	–	–
25	4	06/09/16	223	200	189	–	–
26	4	12/09/16	186	218	198	–	–
27	4	28/09/16	425	185	195	–	–
28	4	20/09/16	198	216	377	–	–
29	4	12/10/16	182	190	160	–	–
30	4	24/10/16	177	171	193	–	–
31	4	07/10/16	266	229	173	–	–
32	4	05/10/16	201	177	169	–	–
33	4	14/10/16	230	194	423	–	–
34	4	19/12/16	360	182	182	–	–
35	4	28/10/16	148	238	355	–	–
36	5	11/01/17	216	149	181	177	–
37	5	22/06/16	175	173	452	173	–
38	5	07/10/16	193	175	175	183	–
39	5	14/09/16	164	171	182	365	–
40	5	01/10/16	184	175	205	182	–
41	5	17/11/16	191	174	167	174	–
42	5	10/10/16	218	192	152	203	–
43	5	12/07/16	140	169	136	318	–
44	5	01/12/16	160	199	182	187	–

Table 1. Cont.

Case	No. of records	Record 1	Record 2	Record 3	Record 4	Record 5	Record 6
45	5	10/08/16	160	157	180	356	–
46	5	01/10/16	157	179	180	313	–
47	5	14/09/16	203	202	226	245	–
48	6	30/07/16	191	214	206	169	169
49	6	13/06/16	263	204	182	160	141
50	6	23/07/16	142	164	184	214	82

No. 008430 of 1993, particularly with title 11 that refers to research in humans. This research was considering without a risk, since calculations were performed on values of clinical diagnostic tests previously indicated by a treating physician. Privacy and integrity of all participants' data were guaranteed throughout the study [17, 18].

Results

A total of 184 records were analyzed for 50 cases, for which 2 cases had 1 record, 7 cases presented 2 records, 8 cases 3 records, 18 cases 4 records, 12 cases 5 records, and 3 cases presented 6 records (Table 1). As for the mathematical induction, it was observed that CD4+ counts differed between 506 and 1,429 cells/μl^3, and the absolute leukocyte counts varied between 3.9 and 16.5 cells/mm^3 (Table 2). These indicated that for cases with leukocyte counts greater than 3.9 cells/mm^3, their CD4+ cell count would be predicted greater than 500 cells/μl^3. Later application of the predictive parameter now established with the remaining 43 cases revealed that 100% of cases, in which absolute white blood cell counts were equal to or greater than 3.9 cells/mm^3 were predicted as CD4+ counts higher than 500 cells/μl^3, thus confirming that the predictive parameter previously established during the inductive process was maintained. Temporal dynamics for the absolute CD4+ counts of a case is shown in Figure 1, while Figure 2 presents these dynamics for the absolute leukocyte counts. Finally, sensitivity and specificity results were obtained for 100% in the blinded study.

Table 2. Total leukocyte and CD4+ counts for 7 prototypical cases of the entire sample

Case		CD4+	Leukocyte count
P9			
	Min.	704	13.4
	Max.	729	16.5
P14			
	Min.	648	5.5
	Max.	918	7.4
P24			
	Min.	549	5.2
	Max.	767	7.6
P25			
	Min.	526	4.1
	Max.	809	5.0
P37			
	Min.	506	3.9
	Max.	687	5.5
P47			
	Min.	742	5.6
	Max.	1,429	7.6
P48			
	Min.	945	5.5
	Max.	1,038	7.7
Total			
	Min.	506	3.9
	Max.	1,429	16.5

Discussion and conclusions

This was the first investigation using the theory of probability, in which predictions were made from leukocyte values equal to or greater than 3.9 cells/mm^3 to find CD4+ lymphocyte counts greater than 500 cells/μl^3. Results of sensitivity and specificity of 100% with a predictive precision were obtained, similarly to 100% results obtained in the blinded study among 50 cases, with several records over time. To reach these results, we first observed the absolute counts of leukocytes and CD4+ lymphocytes from the initial registry of prototypical cases. Subsequently, the counts were evaluated, both for leukocytes and lymphocytes in the second record, which determined if the prediction was correct. Finally, with the observation that, if the third record presented leukocyte values equal to or greater than 3.9 cells/mm^3, it can be deduced that the absolute CD4+ counts would be greater than or equal to 500 cells/μl^3.

Due to the growing interest of making predictions in medicine, particularly in HIV-infected populations, different studies have been developed to predict variables, such as absolute

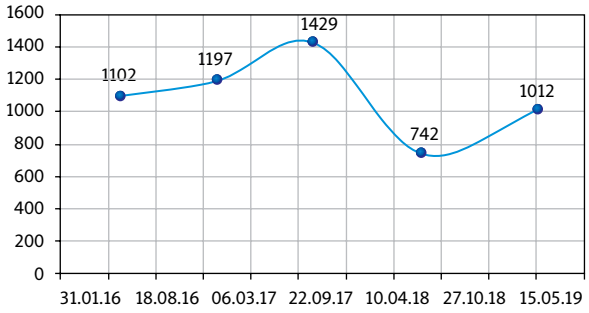

Figure 1. Temporal dynamics of CD4+ counts analyzed in case P47

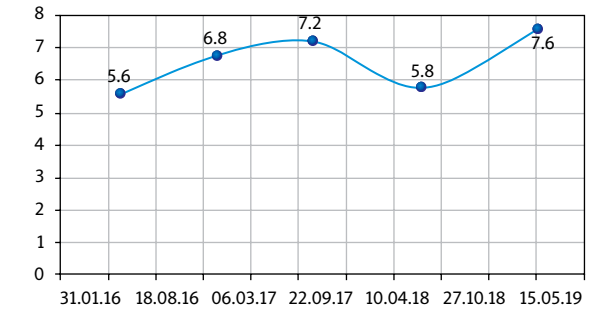

Figure 2. Temporal dynamics of leukocyte counts analyzed in case P47

CD4+ counts based on the absolute white blood cell counts [19-21], ratio CD4/CD8 [22] or CD4+ % [23], and more recently through data mining techniques [24]. Interestingly, it has been found that correct classifications can be made with 99% accuracy using the random forest algorithm for CD4+ counts between 0 and 100, attributing weight to variables, such as age or marital status [24]. However, as these methods are directly related to study populations from the epidemiological point of view, in which they may occasionally require specialized software to carry out the procedures, their applications may be limited. Therefore in this sense, the methodology presented here is simpler when considering only the absolute count of leukocytes to evaluate temporal predictions.

Additionally, it should be noted that methodological foundations of this research are related to the causality of contemporary physics [25], where it is intended to simplify the study phenomenon. That is why the use of other variables related to clinical history, factors of risk, basal treatments, age, among others, are ignored. Despite obtaining 100 percent results, studies should be conducted with larger populations, and where there are different values of CD4+ in clinical interest ranges, such as counts of less than 200 cells/μl^3 or between 200 and 500 cells/μl^3, and to elucidate orders between these ranges.

From this same perspective, different physical theories have been applied to investigate the complexity of biological phenomena that are of interest in medicine, such as obtaining predictions of mortality in intensive care units from hemodynamic variables [26], predictions of malaria epidemics [27], binding of peptides to HLA class II [28], predictions of cardiac dynamics [29, 30], or characterization of bodily structures with fractal geometry [31].

Disclosures

1. Institutional review board statement: Not applicable.
2. Assistance with the article: This is a product derived from the project INV-HUM-3479, funded by the Vice-Rectory of Research of the Universidad Militar Nueva Granada and also, it is derived from the Project 022 financed by Servicios y Asesorías en Infectología, to members of whom we would like to thank for their support of our research. We also thank the Universidad Militar Nueva Granada's Faculty of Medicine for their support to the Deepening Line Course "Physical Theories and Mathematics Applied to Medicine". Also, we extend our gratitude to the Hospital Universitario Nacional de Colombia, especially Doctor Giancarlo Buitrago, Director of the Instituto de Investigaciones Clínicas at Universidad Nacional de Colombia Hospital Universitario Nacional de Colombia and Doctor Rubén Caycedo, Head of the Surgery Department of the Faculty of Medicine at Universidad Nacional de Colombia-Hospital Universitario Nacional de Colombia, for their support.
3. Financial support and sponsorship: Project INV-HUM-3479, funded by the Vice-Rectory of Research of the Universidad Militar Nueva Granada. Project 022 financed by Servicios y Asesorías en Infectología.
4. Conflicts of interest: None.

References

1. Gnedenko BV, Khintchine A. Introducción a la teoría de las probabilidades. Barcelona: Montaner y Simon, S.A.; 1968, p. 175-176.
2. Spiegel M. Estadística, serie Schaum. México: Mc Graw Hill; 1980.
3. Feynman R, Leighton RB, Sands M. Física. Vol. 1. Chapter 6. México: Addison Wesley; 1998.
4. Obregón I. La magia y belleza de las probabilidades. In: Obregón I. Magia y belleza de las matemáticas y algo de su historia. Colombia: Intermedio Ed.; 2007, p. 116.
5. UNAIDS. Global HIV & AIDS statistics – 2018 fact sheet. 2018. Available at: https://www.unaids.org/en/resources/fact-sheet (Accessed: 26.05.2019).
6. De Oliveira M, Bastos M, Martins E, de Lima NA, Leite AJ, Kallas E, et al. Acute HIV infection with rapid progression to AIDS. Braz J Infect Dis 2010; 14: 291-293.
7. Noda AL, Vidal LA, Pérez JE, Cañete R. Interpretación clínica del conteo de linfocitos T CD4 positivos en la infección por VIH. Rev Cubana Med 2013; 52: 118-127.
8. Clift IC. Diagnostic flow cytometry and the AIDS pandemic. Lab Med 2015; 46: e59-e64. DOI: 10.1309/LMKHW2C86ZJDRTFE.
9. Zijenah L, Kadzirange G, Madzime S, Borok M, Mudiwa C, Tobaiwa O, et al. Affordable flow cytometry for enumeration of absolute CD4+ T-lymphocytes to identify subtype C HIV-1 infected adults requiring antiretroviral therapy (ART) and monitoring response to ART in a resource-limited setting. J Transl Med 2006; 14: 33. DOI: 10.1186/1479-5876-4-33.
10. Brown ER, Otieno P, Mbori DA, Farquhar C, Obimbo EM, Nduati R, et al. Comparison of CD4 cell count, viral load, and other markers

for the prediction of mortality among HIV-1-infected Kenyan pregnant women. J Infect Dis 2009; 199: 1292-1300.
11. Azzoni L, Foulkes A, Liu Y, Li X, Johnson M, Smith C, et al. Prioritizing CD4 count monitoring in response to ART in resource-constrained settings: a retrospective application of prediction-based classification. PLoS Med 2012; 9: e1001207. doi: 10.1371/journal.pmed.1001207.
12. Gitura B, Joshi MD, Lule GN, Anzala O. Total lymphocyte count as a surrogate marker for CD4+ T cell count in initiaing antiretroviral therapy at Kenyatta National Hospital, Nairobi. East Afr Med J 2007; 84: 466-472.
13. Sing Y, Mars M. Support vector machines to forecast changes in CD4 count of HIV-1 positive patients. Sci Res Essays 2010; 5: 2384-2390.
14. Foulkes AS, Azzoni L, Li X, Johnson MA, Mounzer K, Montaner LJ. Prediction based classification for longitudinal biomarkers. Ann Appl Stat 2010; 4: 1476-1497.
15. Rodríguez JO, Prieto SE, Correa C, Pérez CE, Mora JT, Bravo J, et al. Predictions of CD4 lymphocytes' count in HIV patients from complete blood count. BMC Med Phys 2013; 13: 3: doi: 10.1186/1756-6649-13-3.
16. Rodríguez J, Prieto S, Correa C, Melo M, Dominguez D, Olarte N, Suárez D, et al. Prediction of $CD4^{+}$ cells counts in HIV/AIDS patients based on sets and probability theories. Curr HIV Res 2018; 16: 416-424.
17. Asociación Médica Mundial. Declaración de Helsinki. Fortaleza, Brasil. 2013. Available at: http://www.isciii.es/ISCIII/es/contenidos/fd-investigacion/fd-evaluacion/fd-evaluacion-etica-investigacion/Declaracion-Helsinki-2013-Esp.pdf (Accessed: 26.05.2019).
18. Ministerio de Salud. Resolución Número 8430 de 1993. Bogotá, Colombia. 1993. Available at: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/RESOLUCION-8430-DE-1993.PDF (Accessed: 26.05.2019).
19. Obirikorang C, Quaye L, Acheampong I. Total lymphocyte count as a surrogate marker for CD4 count in resource-limited settings. BMC Infect Dis 2012; 12: 128. DOI: 10.1186/1471-2334-12-128.
20. Chen J, Li W, Huang X, Guo C, Zou R, Yang Q, et al. Evaluating total lymphocyte count as a surrogate marker for CD4 cell count in the management of HIV-infected patients in resource-limited settings: a study from China. PLoS One 2013; 8: e69704. doi: 10.1371/journal.pone.0069704.
21. Shapiro NI, Karras DJ, Leech SH, Heilpern KL. Absolute lymphocyte count as a predictor of CD4 count. Ann Emerg Med 1998; 32: 323-328.
22. Sauter R, Huang R, Ledergerber B, Battegay M, Bernasconi E, Cavassini M, et al. CD4/CD8 ratio and CD8 counts predict CD4 response in HIV-1-infected drug naive and in patients on cART. Medicine (Baltimore) 2016; 95: e5094. doi: 10.1097/MD.0000000000005094.
23. Kidd PG, Cheng SC, Paxton H, Landay A, Gelman R. Prediction of CD4 count from CD4 percentage: experience from three laboratories. AIDS 1993; 7: 933-940.
24. Kebede M, Zegey 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.
25. Einstein A, Infeld L. La evolución de la física. Barcelona: Salvat; 1986.
26. 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: 209-220.
27. Rodríguez J. Spatio-temporal probabilistic prediction of appearance and duration of malaria outbreaks in municipalities of Colombia. J Phys Conf Ser 2019; 1160: 012018. DOI 10.1088/1742-6596/1160/1/012018.
28. Rodríguez J. Teoría de unión al HLA clase II, teoría de probabilidad combinatoria y entropía aplicadas a secuencias peptídicas. Inmunología 2008; 27: 151-166.
29. Rodríguez J, Prieto SE, Dominguez D, Correa C, Melo M, Pardo J, et al. Application of the chaotic power law to cardiac dynamics in patients with arrhythmias. Rev Fac Med 2014; 62: 539-546.
30. Rodríguez J, Prieto S, Ramírez LJ. A novel heart rate attractor for the prediction of cardiovascular disease. Informatics in Medicine Unlocked 2019; 15: 100174. DOI: https://doi.org/10.1016/j.imu.2019.100174
31. Prieto SE, Rodríguez JO, Correa SC, Soracipa MY. Diagnosis of cervical cells based on fractal and Euclidian geometrical measurements: Intrinsic Geometric Cellular Organization. BMC Med Phys 2014; 14: 2. DOI: 10.1186/1756-6649-14-2.