Medical Research Archives

ESMED
European Society of Medicine

OPEN ACCESS

Published: January 31, 2023

Citation: Zugmaier G., Kerkmann S., et al., 2023. Application of Boolean Algebra for Definition of Myeloid Neoplasms, Medical Research Archives, [online] 11(1). https://doi.org/10.18103/mra.v11i1.3456

DOI:
https://doi.org/10.18103/mra.v11i1.3456

ISSN: 2375-1924

RESEARCH ARTICLE

Application of Boolean Algebra for Definition of Myeloid Neoplasms

*Zugmaier Gerhard[1], Sophie Kerkmann[3] and Franco Locatelli[2]

[1]Department of Haematology, Oncology and Immunology, Philipps University Marburg, Marburg, Germany
[2]Department of Pediatrics, Sapienza, University of Rome, Director Department of Pediatric Hematology and Oncology IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy
[3]Goethe University, Frankfurt, Hessen, Germany

*gerhardz@amgen.com

Abstract

The definition of disease has become increasingly complex in the last 20 years. Especially the definition of tumors of hematopoietic and lymphoid tissues includes multiple parameters such as cytology, histology, immune phenotyping, cytogenetics and molecular genetics. Myeloid neoplasms include an especially high number of permutations of these parameters, making a verbal definition of disease complicated and sometimes confusing. Mathematical concepts have the merit to show what can be proven and what can be calculated. Especially the latter is important for a clear definition of disease. Calculation avoids verbal ambiguity. Boolean algebra also sometimes called "mathematical logic" has been shown to be a useful tool of the definition of acute leukemias. This tool was also applied in this study for the definition of the more complex myeloid neoplasms. The binary number system was used that only includes the numbers 0 and 1. Presence of a parameter was coded by 1, absence by 0. Disease was defined not only by algorithms but also by calculation. This way, myeloid neoplasms could successfully be defined by an algebraic system.

Introduction

Classification is the prerequisite of any medical discipline. Disease must be defined, otherwise neither determination of prognosis nor determination of therapy are possible. Classification requires the collaboration of multiple medical disciplines. The complexity of the topic renders classification by a single expert or small group of experts not feasible anymore. Consensus of all academic groups is necessary for the acceptance of a classification. The World Health Organization (WHO) developed an unanimously accepted classification of tumors of hematopoietic and lymphoid tissues in 2008.[1] The WHO classification is based on the cell lineage which is determined by a myeloid or lymphoid or histiocytic/dendritic origin. The lineages are determined by morphology, cytochemistry, immune phenotype and genetics. Also, the definition of myeloid neoplasms is based on these pillars. This way an algorithm can be developed for each disease entity, enabling to determine clear prognostic and therapeutic criteria. Following the outline above for myeloid neoplasms the following diagnostic procedures are used: Peripheral blood smear, bone marrow aspiration, bone marrow. The material is used for the following evaluations: Cytology, cytochemistry, histology, histochemistry, immune phenotype, cytogenetics, molecular techniques. These evaluations are mandatory and conducted in central laboratories.[1] Since the complexity has increased, the WHO has revised the classification of myeloid neoplasms. Mainly newly discovered genetic abnormalities but also improved standardization of morphological criteria have led to the revised application.[2] The revised classification has led an increased number of diagnostic parameters. Various permutations of diagnostic parameters raise the need of a fast approach to find the correct permutations for the diagnosis of disease. Boolean algebra also known as mathematical logic has been demonstrated to be useful for the definition of acute leukemia. Zugmaier and Locatelli have demonstrated that acute leukemia can be defied in a precise manner by the permutations of Boolean algebra[3,4]. This subset of algebra is characterized by the binary system, which permits the coding of any diagnostic parameter by a binary number. A diagnostic parameter can be coded by either 1 or 0. The former means the parameter is present, the latter means it is absent. In this study we set out, how to define myeloid neoplasms by permutations within Boolean algebra. Since we consider the term "mathematical logic" misleading, we only use the term "Boolean algebra". Boolean Algebras can be defined in multiple ways and are introduced in different ways in the literature.[7] Unfortunately, the Notation of Boolean algebra is not unified, even the arithmetic symbols "+" of addition for "v" and "." of multiplication for "∧" are used. The arithmetic symbol "+" is also used for "either or but not both".[7] To avoid this ambiguity, in computer science the terms "AND" instead of "∧" and "OR" instead of "v" are used. Computer chips are designed by only use of NAND or NOR. In this study we use the symbols of computer science: "NOT", "OR", "AND". For "XNOR" we use the 2-sided

arrow ↔, since the latter is better known to the common readership.

Methods

This study is a systematic review of the definitions of myeloid neoplasms by mathematical concepts. The methods have been described in detail elsewhere.[3] In brief, the binary system is used, which can easily be derived from the decimal system: decimal 0 = binary 0, decimal 1 = binary 1, decimal 2 = binary 10, decimal 3 = binary 11, decimal 4 = binary 100 etc. Present items are coded by the number 1, absent items by the number 0. An item can only have one value at a time, 1 for "present" or 0 for "absent". An item can stand for any parameter such as a proposition, symptom, clinical sign, or laboratory value. The content has no relevance. The values of 2 items A and B can be combined by 4 permutations, see table below.

A	B
1	1
1	0
0	1
0	0

There are 2^4 = 16 permutations possible for a combination of AB assigned to a binary number. The permeations of parameters can be coded by a sequence of binary numbers.[4] From the 16 possible permutations the ones used in this study are listed below.

Permutation I

A	B	AB
1	1	1
1	0	0
0	1	0
0	0	0

A AND B

1 ∧ 1 = 1

1 ∧ 0 = 0

0 ∧ 1 = 0

0 ∧ 0 = 0

Permutation II

A	B	AB
1	1	1
1	0	1
0	1	1
0	0	0

A OR B

This operator was introduced by Jevons in 1864.[5] By doing this, Jevons created today's Boolean algebra, which differs from the one, Boole had created.

1 v 1 = 1
1 v 0 = 1
0 v 1 = 1
0 v 0 = 1

Permutation III

A	B	AB
1	1	1
1	0	0
0	1	0
0	0	1

If and only if A then B

1 ↔ 1 = 1
1 ↔ 0 = 0
0 ↔ 1 = 0
0 ↔ 0 = 1

The algebra is defined as follows:

In the formulas, brackets "()" take precedence over each symbol, "∧" (AND) takes precedence over "v", (OR) the latter takes precedence over "↔".[6]

Distributivity of "v" over "∧": (A v B) ∧ (C v D) = A ∧ C v A ∧ D v B ∧ C v B ∧ D

NOT 1 = 0

NOT 0 = 1

Table 1: Summary of Boolean Operators

A	B	∧ (AND)	V (OR)	↔ (XNOR)	NAND	NOR	XOR		
1	1	1	1	1	0	0	0		
1	0	0	1	0	1	0	1		
0	1	0	1	0	1	0	1		
0	0	0	0	1	1	1	0		

Results

This study is a systemic review of the definitions of myeloid neoplasms by calculation of the diagnosis with the methods of Boolean algebra. For matter of conciseness the diseases are abbreviated as listed below.

Abbreviations

AML acute myeloid leukemia
aCML atypical chronic myeloid leukemia
ANC absolute neutrophil count
AP accelerated phase
BM bone marrow
CML chronic myeloid leukemia
CMML chronic myelomonocytic leukemia
ET essential thrombocythemia
JMML juvenile myelomonocytic leukemia
MDS Myelodysplastic Syndrome
MDS/MPN-RS-T myelodysplastic/myeloproliferative neoplasm with ring sideroblasts and thrombocytosis
MDS/MPN myelodysplastic/myeloproliferative neoplasm, unclassifiable
MDS-U myelodysplastic syndrome, unclassifiable
NOS not otherwise specified
PB peripheral blood
PMF primary myelofibrosis
PV polycythemia vera
RS ring sideroblast

Myeloid neoplasms ↔ myeloproliferative neoplasms OR mastocytosis OR myeloid/lymphoid neoplasms AND eosinophilia AND gene rearrangement OR myelodysplastic/myeloproliferative neoplasms OR myelodysplastic syndrome OR myeloid neoplasm with germline predisposition OR acute myeloid leukemia.

This means in verbal terms that, if the hypothesis of myeloid neoplasm is raised, the diagnosis can be made as soon as at least one of the parameters listed above on the right side of the Boolean equation has the value 1. Then also "myeloid neoplasm" has the value 1.

1) Myeloproliferative neoplasms ↔ chronic myeloid leukemia AND BCR-ABL1 OR chronic neutrophilic leukemia OR polycythemia vera OR primary myelofibrosis OR essential thrombocythemia OR chronic eosinophilic leukemia NOS OR unclassifiable myeloproliferative neoplasm

   1.1) Chronic neutrophilic leukemia ↔ $\geq$25 x 10^9/L PB leukocytes AND $\geq$80% PB neutrophils AND <10% PB neutrophil precursors AND < 10^9/L PB monocytes AND NOT dysgranulopoiesis AND bone marrow cells ↓ AND <5% BM

myeloblasts AND (NOT (BCR-ABL1 OR polycythemia vera OR thrombocythemia OR primary myelofibrosis OR PDGFR (A OR B) OR FGFR1 OR PCM1-JAK2)) AND (C3F3R mutation OR neutrophils ≥3months↑ AND spleen↑)

1.2) ET ↔ ≥ 450 x 10^9/L PB thrombocytes AND BM megakaryocytes ↑ AND NOT (BCR-ABL1 OR PV OR PMF) AND (JAK2 OR CALR OR MPL OR clonal maker OR NOT reactive thrombocytosis).

In this case one parameter is not sufficient for the diagnosis. All parameters connected by AND need to have the value 1. Then also "ET" has the value 1.

2) Mastocytosis ↔ cutaneous mastocytosis OR systemic mastocytosis OR mast cell sarcoma

3) Myeloid/lymphoid neoplasms AND eosinophilia AND gene rearrangement ↔ PDGFRA rearrangement (A OR B) OR FGFR1 OR PCM 1–JAK2

4) Myelodysplastic/myeloproliferative neoplasms ↔ chronic myelomonocytic leukemia OR atypical chronic myeloid leukemia AND NOT BCR-ABL1 OR juvenile myelomonocytic leukemia OR myelodysplastic/myeloproliferative neoplasm AND ring sideroblasts AND (thrombocytosis OR myelodysplastic / myeloproliferative neoplasm unclassifiable)

4.1) CMML ↔ ≥10^9/L AND ≥10% PB monocytes AND NOT (CML OR PMF OR ET) AND NOT ((PDGFR (A OR B) OR PDGFR1 OR PCM1-JAJ2)) AND < 20% PB AND BM blasts AND (≥1 lineage dysplasia OR genetic abnormality OR ≥3 months monocytes↑)

4.2) aCML ↔ ≥13x10^9/L PB leukocytes AND ≥10% neutrophil precursors AND dysgranulopoiesis AND <2% PB basophils AND <10% PB monocytes AND bone marrow cells↑ AND <20% BM blasts AND NOT ((PDGFR (A OR B) OR PCM1-JAK2)) AND NOT (CML OR PMF OR PV OR ET)

4.3) JMML ↔ ≥10^9/L PB monocytes AND (<20% PB AND BM blasts) AND spleen↑ AND NOT BCR-ABL1 AND ((PTNP11 OR (K OR N) RAS OR neurofibromatosis type1 OR NF1 OR loss heterozygosity CBL OR monosomy7 OR ≥2 of (HBF↑; myeloid precursors; erythroid precursors; OR CSF2 hypersensitivity in colony assay; STAT5 phosphorylation↑)

4.4) MDS/MPN-RS-T ↔ HB↓ AND ≥15% RS <1% PB blasts AND <5% BM blasts AND ≥ 450x10^9/L thrombocytes AND (SFB3 OR NOT growth factor therapy) AND NOT ((BCR-ABL1 OR PDGFR (A OR B) OR FGFR1 OR PCM1-JAK2 OR t(3.3)(q21.3; q26.2) OR INV(3)(q21.3q26.2) OR del(5q))

4.5) MDS/MPN-U ↔ (<20% PB AND BM blasts) AND (≥450x10^9/L PB thrombocytes OR ≥450x10^9/L PB leukocytes) AND NOT ((growth factor therapy OR PDGFR (A OR B) OR FGFR1 OR PCM1-JAK2))

5) MDS ↔ single lineage dysplasia OR ring sideroblasts OR multilineage dysplasia OR excess blasts OR del (q5) OR unclassifiable OR refractory cytopenia of childhood

5.1) Cytopenia ↔ <10g/dL HB OR <100x10^9/L thrombocytes OR <1.8x10^9/L ANC

5.2) MDS-U ↔ 1% PB blasts OR pancytopenia OR MDS defining genetic abnormality

6) Myeloid neoplasms AND germline predisposition ↔ AML AND CEBPA mutation OR DDX41 mutation OR RUNX1 mutation OR ANKRD26 mutation OR GATA2 mutation

Discussion

Although disease can in most case be diagnosed in individual patients, the principal definition frequently proves to be challenging. Especially when the number of diagnostic parameters is high, finding the correct permeation can be difficult. For this reason we have followed the approach to implement more mathematics into medicine. In this study, we used Boolean algebra to define a highly complex disease entity such as myeloid neoplasm.

Mathematics has become an established discipline in medicine as it has been for a long time in physics, chemistry and biology.[8] Without biostatistics and mathematical modeling clinical research would not be possible. Boolean algebra as part of discrete mathematics has become an important component of theoretical medicine not only in bioinformatics but also in numerous other fields, such as gene regulatory networks or signaling networks.[9-11] Biochemical reaction networks, Boolean models of gene regulatory networks, algebraic statistic in genomics provide further examples for the use of mathematics in biology.[12] Boolean algebra has been demonstrated to be useful detect the association of expression levels in gene models with the outcome of disease.[13] A NOT- gated signal integrator called the T mod system was developed for patients with malignant solid tumors. Together with strict selection for defined lesions in the tumor genomes this approach has enabled selective therapies for certain patients with cancer.[14] Boolean models have been used in pathology[15] for prediction of clinical outcome in acute myeloid leukemia[16] and acute leukemias[4].

Definition of disease by Boolean algebra is only possible, if the connections of parameters are clearly defined. This is not always the case. In myeloid neoplasms Boolean algebra could not be used in some cases. Although myelodysplastic syndrome comprises of many detailed criteria of diagnosis, their combinations are not consistent. In table 6.01 of The WHO criteria of 2016 one criterion of the division of

myelodysplastic syndrome with excess blasts includes the following permutations of parameters: Bone marrow blasts 5 -19% or peripheral blood last 2-4%, bone marrow blasts <10% and peripheral blood last < 5%, no Auer rods. It's not clear, whether these parameters are exclusive or inclusive and how they need to be combined. In Table 2.03 of the WHO classification of Tumors of the Hematopoietic and Lymphoid Tissues polycythemia vera is defined by 3 major criteria and 1 minor criterion. The diagnosis is met, if all 3 major criteria or the first 2 major criteria and the 1 minor criterion are present. The correct definition would have been the to require the first 2 criteria and at least one of the 2 remaining criteria. The terms "major" and "minor" lead to confusion in this case.

One of the problems of Boolean algebra is the lack of standardization of the rotation. The symbols for the Boolean operators vary. Especially between Europe and US, there is no consensus on the selection of symbols.[7]Another issue is caused by using the same symbols in different settings of mathematics. For these reasons, we use the acronyms of computer science in this study.
We have decided to use the term "Boolean algebra" and avoid the term "mathematical logic". It's not our intention to dismiss logic. Logic is part of mathematics as it is part of any other science and any rational endeavor in general.[17] However, some authors pool logic and mathematics with a mix of mathematical and verbal concepts, which define each other mutually in a circular way.[18] In order to eliminate all "logic elements", "truth values" were replaced by numbers and. This way confusing "logic operations" were avoided and only arithmetic operations were applied.[19] In this study we have used the binary system, which has the same structure as the decimal system. This structure is important, since all measurable parameters are defined by natural numbers. Other systems applied in the biological science such as p-adic number systems use a different approach. They do not base concepts on numbers but vice versa base numbers on geometrical concepts.[20] For this reason, the p-adic number system would result in a step backwards if applied in medicine.

In conclusion, the use of Boolean algebra has shown to be a promising approach to make the diagnosis of disease calculable. The mathematical approach applied in this study does not have the ambition to clarify the nature of disease or pathophysiologic mechanisms. Mathematical methods can relieve the busy clinician from mechanical tasks, which may prevent her or him from interacting with the patient

However, the methods described in this study will never replace the physician, who will always have to make the last decision have the last word on the best methods to serve patients.

Corresponding author:
Zugmaier Gerhard
Department of Haematology, Oncology and Immunology, Philipps University Marburg, Marburg, Germany
Email: gerhardz@amgen.com

Disclosure Statement
None

Conflict of Interest
None

Funding
None

Conflicts of interest:
None

References:

1. Swerdlow, SH et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*, World Health Organization Classification of Tumours. 4th ed. Lyon: International Agency for Research on Cancer; 2008.

2. Swerdlow SH et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Revised 4th ed. Lyon: International Agency for Research on Cancer; 2017.

3. Zugmaier G, Locatelli F. Application of Mathematical Logic for Immunophenotyping of B-Cell Precursor Acute Lymphoblastic Leukemia (BCP-ALL). *Biomedical Genetics and Genomics*. 2019;4: 1-3. doi: 10.15761/bgg.1000148

4. Zugmaier G, Locatelli F. Application of Mathematical Logic for Cytogenetic Definition and Risk Stratification of B-Cell Precursor Acute Lymphoblastic Leukemia (BCP-ALL). *Medical Research Archives*. 2021;9(2):1-8. doi: 10.18103/mra.v9i2.2328

5. Brown FM. *Boolean Reasoning: The Logic of Boolean Equations*. 2nd ed. Mineola, New York: Dover; 2003.

6. Takeuti G. *Proof Theory*. 2nd ed. Amsterdam, North Holland: Dover Publications; 1987.

7. Hoffmann DW. *Grundlagen der technischen Informatik*. 5. Auflage. München: Carl Hanser Verlag; 2016

8. Matthäus F, Matthäus S, Harris S, Hillen T. *The Art of Theoretical Biology*. Springer; 2020.

9. Albert R, Robeva R. Signaling Networks: Asynchronous Boolean Models. *Algebraic and Discrete Mathematical Methods for Modern Biology*. 2015:65-91. doi: 10.1016/b978-0-12-801213-0.00004-6

10. He Q, Macauley M, Davies R. Dynamics of Complex Boolean Networks. *Algebraic and Discrete Mathematical Methods for Modern Biology*. 2015:93-119. doi: 10.1016/b978-0-12-801213-0.00005-8

11. Lin P-CK, Khatri SP. *Logic Synthesis for Genetic Diseases: Modeling Disease Behavior Using Boolean Networks*. New York, New York: Springer; 2014.

12. Macauley M, Youngs N. The Case for Algebraic Biology: From Research to Education. *Bulletin of Mathematical Biology*. 2020;82(9):115. doi: 10.1007/s11538-020-00789-w

13. Varadan V, Anastassiou D. Inference of Disease-Related Molecular Logic from Systems-Based Microarray Analysis. *PLoS Computational Biology*. 2006;2(6): 585-597. doi: 10.1371/journal.pcbi.0020068.eor

14. DiAndreth B, Hamburger AE, Xu H, Kamb A. The TMOD Cellular Logic Gate as a Solution for Tumor-Selective Immunotherapy. *Clinical Immunology*. 2022;241: 1-8. doi: 10.1016/j.clim.2022.109030

15. Riede U, Moore GW, Williams MB. Quantitative Pathology by Means of Symbolic Logic. *CRC Critical Reviews in Toxicology*. 1983;11(4):279-332. doi: 10.3109/10408448309037457

16. Palma A, Iannuccelli M, Rozzo I, et al. Integrating Patient-Specific Information into Logic Models of Complex Diseases: Application to Acute Myeloid Leukemia. *Journal of Personalized Medicine*. 2021;11(2):117. doi: 10.3390/jpm11020117

17. Grattan-Guinness I. Mathematics and Symbolic Logics: Some Notes on an Uneasy Relationship. *History and Philosophy of Logic*. 1999;20(3-4):159-167. doi: 10.1080/01445349950044116

18. Leitgeb H. Hype: A System of Hyperintensional Logic (With an Application to Semantic Paradoxes). *Journal of Philosophical Logic*. 2019;48(2):305-405. doi: 10.1007/s10992-018-9467-0

19. Steffens HJ, Muehlmann K, Zoellner C. *Mathematik für Informatiker für Dummies*. Weinheim: Wiley – VCH; 2020

20. Hua H, Hovestadt L. P-adic Numbers Encode Complex Networks. *Nature*. 2021;11(17). doi: 10.1038/s41598-020-79507-4