APPLICABILITY OF LOTKA'S LAW TO GLOBAL OPEN ACCESS RESEARCH ON POLYCYSTIC OVARY SYNDROME: A SCIENTOMETRIC STUDY

Asmat Ali, (PhD, Research scholar, Annamalia University, Department of Library and Information Sciences, Annamalainagar- Tamilnadu –asmatali999@gmail.com

Dr. R. Ramesh, (Assistant professor, Department of Library and Information sciences, Annamalia University Annamalainagar- Tamilnadu)

Nahida Hassan, (Masters in Statistics, Islamic University of Science and Technology- Awantipora,J&K)

ABSTRACT :
Polycystic Ovary Syndrome (PCOS) is multifactorial endocrinal disorder involving a complex interplay of genetic, environmental, and hormonal factors. The present article is an attempt to trace the open access scholarly output in the field of Polycystic Ovary Syndrome (PCOS) at global level. For achieving the set objectives the publication data collected from the Web of Science to the time span of 2020-2023 was systematically analyzed and organized to meet the study's objectives using bibliometric indicators like Document Type, open Access type, prolific Authors, Prolific Journals, and Keywords. This study has applied Lotka's law pertaining to authorship productivity within the domain of PCOS literature therein; A Kolmogorov-Smirnov (K-S) test was employed to evaluate the concordance between the distributions of the empirical data set. The findings indicate that (73.802%) publications pertaining to PCOS have been disseminated through the Green mode of open access. The maximum frequency of publications on PCOS is found in the form of Article representing 2790 publications (67%).It has been ascertained that China is the predominant nation in disseminating research on Polycystic Ovary Syndrome within open access platform. Shandong University in China emerges as the preeminent institution, contributing (258, 6.21%) articles. furthermore, Kolmogorov Simonov test is applied for the wellness of the Lotka's law for the estimations of Lotka's types acquired from least square methods and watched authorship information distribution holds good for the Lotka's law.
Key-words:
Polycystic Ovary Syndrome, Open Access, PCOS, Scientometrics, Lotka's law,, Kolmogorov- Smirnov Test, and VOSviewer.

INTRODUCTION:
PCOS is multifactorial endocrinal disorder, involving a complex interplay of genetic, environmental, and hormonal factors. Family studies indicate a strong genetic component, with multiple genes implicated in the pathogenesis, including those related to insulin resistance and androgen biosynthesis[1]. PCOS diagnosis relies on the Rotterdam criteria (2003), which require two out of the following three features for diagnosis: oligo- or anovulation, clinical or biochemical signs of hyperandrogenism, and polycystic ovaries on ultrasound. However, there remains some controversy regarding these criteria, particularly the inclusion of polycystic ovarian morphology, as it may be present in women without PCOS[2]. The present study of PCOS using the Scientometric method will likely provide further insights and improve outcomes for women affected by this condition. This study both adds to and supports empirical research on PCOS literature and offers insightful information about the trajectory of earlier research.

DATA COLLECTION:

For the present study, the publication data was retrieved and downloaded from the Web of Science. The advanced search strategy for Global Research output was formulated; the search string used for data extraction was: TS= ("Polycystic Ovarian Syndrome" OR ""), refined by: ALL OPEN ACCESS and time span 2020-2023. A total of 4153 records were retrieved using the search query. The data as such retrieved was analysed and tabulated in an efficient manner to gauge the findings in order to achieve the set objectives. A few tools, including Bib Excel Software for citation analysis and Microsoft Excel for data visualization, were utilized to analyze the collected documents .Further, VOSviewer, a mapping tool was used. The data was downloaded on Aug 23, 2024.

OBJECTIVES:

1. To analyse the document type and Open access type publications on Polycystic Ovary Syndrome (PCOS) literature during the period of 2020-2023.
2. To identify prolific authors in the field of Polycystic Ovary Syndrome (PCOS) through open access publications at a global level.
3. To Test the Applicability of Lotka's Law in the author productivity
4. To determine whether the "n" value confirms to Lotka's Law through K- S test.

LITERATURE REVIEW:

[3]An empirical examination of Lotka's law alongside the patterns of author productivity within the domain of Artificial Intelligence was conducted. K-S statistical test was employed to evaluate the level of concordance between the observed data distribution and the theoretical inverse power relationship, specifically with the theoretical parameter set as a=2. The findings indicated that Lotka's inverse square law is indeed adhered to in this context.

The prevailing status of the scholarly literature pertaining to social network analysis (SNA) was analyzed by using bibliometric analysis and found that the United States is leading in terms of published papers and `Computers in Human Behavior` is the major journal in which SNA-related articles were disseminated. The application of the Kolmogorov–Smirnov (K-S) test reflecting author productivity distribution adhered to Lotka's law[4]

[5]An analysis of ten open access journals in the field of Library and Information Science (LIS) was carried utilizing bibliometric indicators and revealed that the prevalence of single authorship is predominant within these LIS open access publications. The calculated citation indices, including the Citation Index (CI), Degree of Collaboration (DC), and Collaboration Coefficient (CC), were determined to be 1.95, 0.47, and 0.29, respectively. The findings do not conform to Lotka's inverse square law.

[6]A comparative analysis of data downloaded from the Web of Science (WoS) of time span 2000 to 2020. Lotka's law pertaining to author productivity alongside the Google Scholar i10-Index within the discipline of chemistry in Iran was employed. The most prolific Iranian contributors found from the analysis were Mohamadreza Ganjali from the University of Tehran, Majid Heravi from Alzahra University, and Mojtaba Shamsipur from the Razi University of Kermanshah. Furthermore, the empirical validation of Lotka's law was not established in the assessment of productivity among Iranian authors in the realm of chemistry.

DOCUMENT SOURCE TYPE:

Figure- 1 depicts the document source types in which the Open Access scientific literature is published globally on polycystic ovary syndrome from 2020-2023.The maximum frequency of publications on

PCOS is found in the form of Article representing 2790 publications (67%) whereas review is found second representing 1072 publications (26%) and editorial material 90 (2.0%) ranking at third.
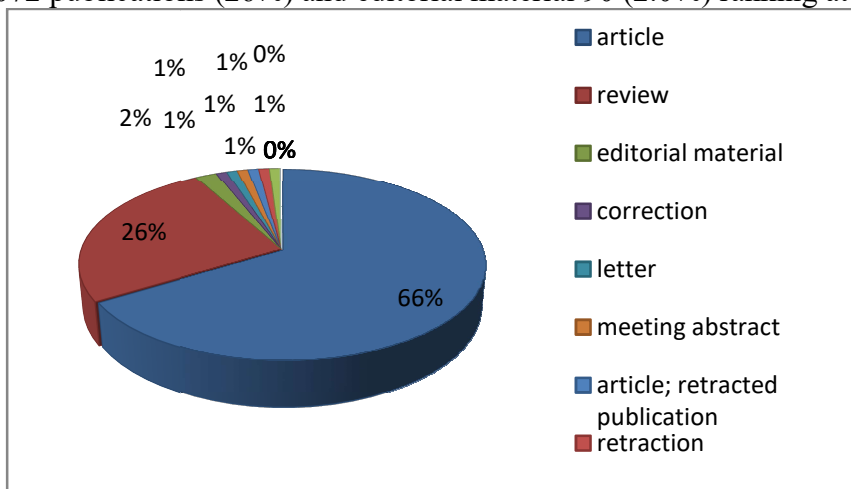

Figure-1: Document source Type

Open access modes of publications on PCOS:

Table 1: Open access Type		
Open Access	Publications	% of 4,153
All Open Access	4153	100
Gold	2805	67.542
Gold-Hybrid	494	11.895
Free to Read	520	12.521
Green Published	3065	73.802
Green Accepted	222	5.346
Green Submitted	240	5.779

Table-1 presents a comprehensive overview of the various models for open-access publication, along with their corresponding frequencies. It has been determined that a total of 3065 (73.802%) publications pertaining to PCOS have been disseminated through the Green mode of open access, thereby securing the first position, followed by the Gold mode of open access, which ranks second with 2805 publications (67.542%), and the free to read mode of open access, which occupies the third position with 520 publications (12.521%).

Prolific Authors on open access publications on PCOS:

Table:2 Prolific Authors			
Authors	**Country**	**Articles**	**%age**
Wang ,Y	China	133	3.20
Zhang, Y	China	124	2.98
Liu ,Y	China	99	2.38
Li ,Y	China	81	1.95
Chen ,Y	China	73	1.75
Zhang, J	China	71	1.70
Li ,J	China	66	1.58

Wang, L	China	64	1.54
Li X,	China	63	1.51
Wang	China	61	1.46

The list of ten top authors who produced the highest contribution to research output on open access publications of polycystic ovary syndrome at global level are all from china. of Wang ,Y the most productive author with 133(3.20%) publications followed by Zhang, Y 124(2.98%) and Liu,Y 99(2.38%) .It is also noted that 2 out of 10 prolific authors contributed more than a hundred research publications each while the rest eight authors contributed more than 60 publications each.
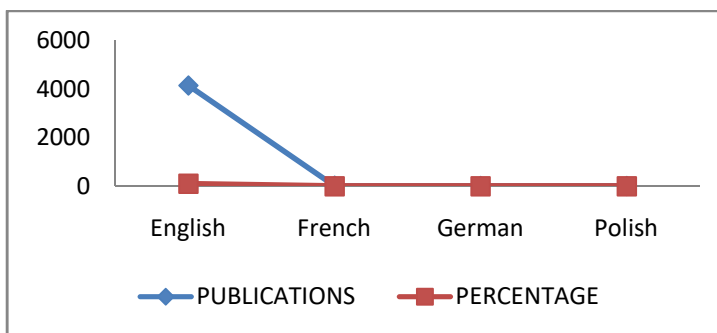
Language wise publications

Figure-2: Language wise publications on open access PCOS publications

The analysis of data presented in Figure-2 shows 4153 papers published in 4 languages. It can be seen from figure -2 that English has been used as a significant communication language for open access polycystic ovary syndrome publications. Nearly 99.80% of publications appear in the English language and dominates in the first place out of four languages, followed by French (5.744%), German (0.751%) and polish (0.140%).

INSTITUTIONAL OUTPUT:

The impact of scholarly output on global open access can be elucidated from an institutional standpoint. As illustrated in Figure - 3, Shandong University in China emerges as the preeminent institution, contributing (258, 6.21%) articles, succeeded by Shanghai Jiao Tong University from China (254, 6.11%), and Zhejiang University, which occupies the third position with (215, 5.17%). It has been ascertained that China is the predominant nation in disseminating research on Polycystic Ovary Syndrome within open access platform.
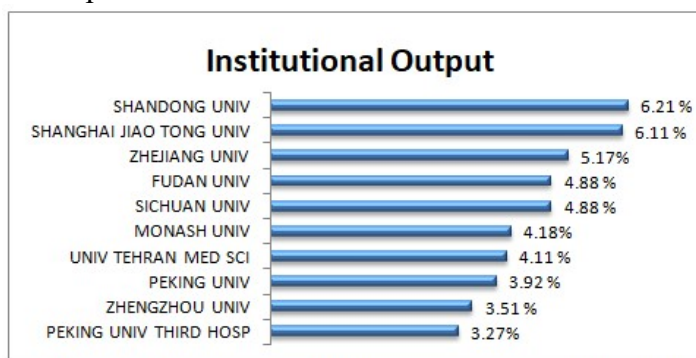

Figure-3: Language wise publications on open access PCOS publications

ANALYSIS OF LOTKA'S LAW:-

Table:3 Lotka's Law on PCOS (Open Access)				
x	**y**	**X = Logx**	**Y = Logy**	**XY**	**X²**
1	13207	0.0000	4.1208	0.0000	0.0000
2	1983	0.3010	3.2973	0.9926	0.0906
3	715	0.4771	2.8543	1.3619	0.2276
4	340	0.6021	2.5315	1.5241	0.3625
5	198	0.6990	2.2967	1.6053	0.4886
6	152	0.7782	2.1818	1.6978	0.6055
7	95	0.8451	1.9777	1.6714	0.7142
8	57	0.9031	1.7559	1.5857	0.8156
9	45	0.9542	1.6532	1.5776	0.9106
10	27	1.0000	1.4314	1.4314	1.0000
11	30	1.0414	1.4771	1.5383	1.0845
12	31	1.0792	1.4914	1.6094	1.1646
13	19	1.1139	1.2788	1.4245	1.2409
14	10	1.1461	1.0000	1.1461	1.3136
15	12	1.1761	1.0792	1.2692	1.3832
16	13	1.2041	1.1139	1.3413	1.4499
17	10	1.2304	1.0000	1.2304	1.5140
18	10	1.2553	1.0000	1.2553	1.5757
19	12	1.2788	1.0792	1.3800	1.6352
20	8	1.3010	0.9031	1.1749	1.6927
21	8	1.3222	0.9031	1.1941	1.7483
22	5	1.3424	0.6990	0.9383	1.8021
23	2	1.3617	0.3010	0.4099	1.8543
24	4	1.3802	0.6021	0.8310	1.9050
25	5	1.3979	0.6990	0.9771	1.9542
26	5	1.4150	0.6990	0.9890	2.0021
27	5	1.4314	0.6990	1.0005	2.0488
28	2	1.4472	0.3010	0.4356	2.0943
29	1	1.4624	0.0000	0.0000	2.1386
30	1	1.4771	0.0000	0.0000	2.1819
31	7	1.4914	0.8451	1.2603	2.2242
32	5	1.5051	0.6990	1.0521	2.2655
33	2	1.5185	0.3010	0.4571	2.3059

34	2	1.5315	0.3010	0.4610	2.3454
35	3	1.5441	0.4771	0.7367	2.3841
36	2	1.5563	0.3010	0.4685	2.4221
37	2	1.5682	0.3010	0.4721	2.4593
38	2	1.5798	0.3010	0.4756	2.4957
39	1	1.5911	0.0000	0.0000	2.5315
41	1	1.6128	0.0000	0.0000	2.6011
42	2	1.6232	0.3010	0.4886	2.6349
43	4	1.6335	0.6021	0.9834	2.6682
47	2	1.6721	0.3010	0.5034	2.7959
48	1	1.6812	0.0000	0.0000	2.8266
52	1	1.7160	0.0000	0.0000	2.9447
55	1	1.7404	0.0000	0.0000	3.0289
61	1	1.7853	0.0000	0.0000	3.1874
63	1	1.7993	0.0000	0.0000	3.2376
64	1	1.8062	0.0000	0.0000	3.2623
66	1	1.8195	0.0000	0.0000	3.3107
71	1	1.8513	0.0000	0.0000	3.4272
73	1	1.8633	0.0000	0.0000	3.4720
81	1	1.9085	0.0000	0.0000	3.6423
99	1	1.9956	0.0000	0.0000	3.9826
124	1	2.0934	0.0000	0.0000	4.3824
133	1	2.1239	0.0000	0.0000	4.5107
	17060	**77.0352**	**45.1568**	**40.9515**	**116.3482**

To achieve the objective of testing the applicability of Lotka's law of author production in open access scholarly publications on PCOS research, the analysis of Lotka's law is carried out and it is simplified in tables

Consequently, the first step is to calculate the exponent n using the least-squares method and according to the following formula:

$$n = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

All the data needed for the n formula can be obtained from Table -3 **.**The only index that requires further work is N, which represents the number of pairs considered. In this example, those authors who have published between one and 133 articles will be considered, representing 56 pairs of data ($N = 56$).

$$n = \frac{56(40.9515) - (77.0352)(45.1568)}{56(116.3482) - (77.0352)^2}$$

$$□ = \frac{1185.3791}{581.0772}$$
$$□ = 2.04$$

Thus, the value of *n* (absolute value) is 2.04, which will then be the specific value of the coefficient in Lotka's formula that will explain author productivity in this particular case.

The theoretical value of '*n*' =2.04 is matched with the table value of R. Rosseau for getting C.S. value 0.6215

Constant Value of Present Study	n Value
0.6215	2.04
Lotka's Constant Value	n Value
0.6079	2

D-Max Value Present Study　　　　　　　　D-Max Value of Lotka's Study
　　　0.15265　　　　　　　　　　　　　　　　　0.016625

Finally, the Kolmogorov-Smirnov test is applied to verify whether the observed data fit the theoretical distribution according to Lotka's law. The highest value in column (D_{max}) is taken as reference for comparison with the critical value (c.v.), whose general formulation is:

$$□_{□□□} = □(□) - □□(□)$$
$$\hat{□} = 2.04$$

Theoretical Value of C = 0.6215

$$□□^{+} = 0.6215 \left(\frac{1}{□^{2.04}}\right)$$

$D_{MAX} = 0.15265$

Critical Value at 0.01 level of significance = $\frac{2.04}{\sqrt{17060}} = 0.0156$

KOLMOGOROV- SMIRNOV TEST:

Table: 4 Kolmogorov- Smirnov Test									
x	**y**	Observe d = y/Σyx	Value = Σ(y/Σyx)	Expected Frequency	Value of Frequency /Cumulative	Difference (D)	Expected Frequency	Value of Frequency / Cum.	Diff
1	13207	0.77415	0.77415	0.62150	0.62150	0.15265	0.60790	0.60790	0.16625
2	1983	0.11624	0.89039	0.15113	0.77263	-0.03489	0.14782	0.75572	-0.03158
3	715	0.04191	0.93230	0.06609	0.83871	-0.02418	0.06464	0.82036	-0.02273
4	340	0.01993	0.95223	0.03675	0.87546	-0.01682	0.03594	0.85630	-0.01601
5	198	0.01161	0.96383	0.02331	0.89877	-0.01170	0.02280	0.87910	-0.01119
6	152	0.00891	0.97274	0.01607	0.91484	-0.00716	0.01572	0.89482	-0.00681

7	95	0.00557	0.97831	0.01173	0.92658	-0.00617	0.01148	0.90630	-0.00591
8	57	0.00334	0.98165	0.00894	0.93551	-0.00559	0.00874	0.91504	-0.00540
9	45	0.00264	0.98429	0.00703	0.94254	-0.00439	0.00687	0.92191	-0.00424
10	27	0.00158	0.98587	0.00567	0.94821	-0.00409	0.00554	0.92746	-0.00396
11	30	0.00176	0.98763	0.00467	0.95287	-0.00291	0.00456	0.93202	-0.00281
12	31	0.00182	0.98945	0.00391	0.95678	-0.00209	0.00382	0.93584	-0.00200
13	19	0.00111	0.99056	0.00332	0.96010	-0.00221	0.00325	0.93909	-0.00213
14	10	0.00059	0.99115	0.00285	0.96295	-0.00227	0.00279	0.94188	-0.00220
15	12	0.00070	0.99185	0.00248	0.96543	-0.00178	0.00242	0.94431	-0.00172
16	13	0.00076	0.99261	0.00217	0.96760	-0.00141	0.00213	0.94643	-0.00136
17	10	0.00059	0.99320	0.00192	0.96952	-0.00133	0.00188	0.94831	-0.00129
18	10	0.00059	0.99379	0.00171	0.97123	-0.00112	0.00167	0.94998	-0.00109
19	12	0.00070	0.99449	0.00153	0.97276	-0.00083	0.00150	0.95148	-0.00079
20	8	0.00047	0.99496	0.00138	0.97414	-0.00091	0.00135	0.95283	-0.00088
21	8	0.00047	0.99543	0.00125	0.97539	-0.00078	0.00122	0.95405	-0.00075
22	5	0.00029	0.99572	0.00113	0.97652	-0.00084	0.00111	0.95516	-0.00082
23	2	0.00012	0.99584	0.00104	0.97756	-0.00092	0.00101	0.95617	-0.00090
24	4	0.00023	0.99607	0.00095	0.97851	-0.00072	0.00093	0.95710	-0.00069
25	5	0.00029	0.99637	0.00087	0.97939	-0.00058	0.00086	0.95795	-0.00056
26	5	0.00029	0.99666	0.00081	0.98019	-0.00051	0.00079	0.95874	-0.00050
27	5	0.00029	0.99695	0.00075	0.98094	-0.00045	0.00073	0.95947	-0.00044
28	2	0.00012	0.99707	0.00069	0.98163	-0.00058	0.00068	0.96015	-0.00056
29	1	0.00006	0.99713	0.00065	0.98228	-0.00059	0.00063	0.96078	-0.00057
30	1	0.00006	0.99719	0.00060	0.98288	-0.00054	0.00059	0.96137	-0.00053
31	7	0.00041	0.99760	0.00056	0.98345	-0.00015	0.00055	0.96193	-0.00014
32	5	0.00029	0.99789	0.00053	0.98397	-0.00024	0.00052	0.96244	-0.00022
33	2	0.00012	0.99801	0.00050	0.98447	-0.00038	0.00049	0.96293	-0.00037
34	2	0.00012	0.99812	0.00047	0.98494	-0.00035	0.00046	0.96338	-0.00034
35	3	0.00018	0.99830	0.00044	0.98538	-0.00026	0.00043	0.96381	-0.00025
36	2	0.00012	0.99842	0.00042	0.98579	-0.00030	0.00041	0.96422	-0.00029
37	2	0.00012	0.99853	0.00039	0.98619	-0.00028	0.00038	0.96461	-0.00027
38	2	0.00012	0.99865	0.00037	0.98656	-0.00025	0.00036	0.96497	-0.00025
39	1	0.00006	0.99871	0.00035	0.98691	-0.00029	0.00035	0.96531	-0.00029
41	1	0.00006	0.99877	0.00032	0.98723	-0.00026	0.00031	0.96563	-0.00025
42	2	0.00012	0.99889	0.00030	0.98753	-0.00019	0.00030	0.96592	-0.00018
43	4	0.00023	0.99912	0.00029	0.98782	-0.00005	0.00028	0.96621	-0.00005
47	2	0.00012	0.99924	0.00024	0.98806	-0.00012	0.00024	0.96644	-0.00012
48	1	0.00006	0.99930	0.00023	0.98829	-0.00017	0.00023	0.96667	-0.00017
52	1	0.00006	0.99936	0.00020	0.98849	-0.00014	0.00019	0.96686	-0.00013

55	1	0.00006	0.99941	0.00018	0.98867	-0.00012	0.00017	0.96703	-0.00011
61	1	0.00006	0.99947	0.00014	0.98881	-0.00008	0.00014	0.96717	-0.00008
63	1	0.00006	0.99953	0.00013	0.98894	-0.00007	0.00013	0.96730	-0.00007
64	1	0.00006	0.99959	0.00013	0.98907	-0.00007	0.00013	0.96742	-0.00007
66	1	0.00006	0.99965	0.00012	0.98919	-0.00006	0.00012	0.96754	-0.00006
71	1	0.00006	0.99971	0.00010	0.98929	-0.00005	0.00010	0.96764	-0.00004
73	1	0.00006	0.99977	0.00010	0.98939	-0.00004	0.00010	0.96774	-0.00004
81	1	0.00006	0.99982	0.00008	0.98947	-0.00002	0.00008	0.96782	-0.00002
99	1	0.00006	0.99988	0.00005	0.98952	0.00001	0.00005	0.96787	0.00001
124	1	0.00006	0.99994	0.00003	0.98956	0.00003	0.00003	0.96790	0.00003
133	1	0.00006	1.00000	0.00003	0.98959	0.00003	0.00003	0.96793	0.00003
	17060		**Present study D.Max = 0.15265**				**Lotka's D Max = 0.16625**		

The theoretical value of C as 0.0156 for $n = 2.04$ is taken from the book 'Power Laws in the Information Production Process: Lotkanian Informetrics' by Egghe (2005). Kolmogorov Simonov test is applied for the wellness of the Lotka's law for the estimations of Lotka's types acquired from least square methods. The outcomes tabulated in the above table show that the estimation of D-max, i.e., 0.0156 decided with Lotka's type, i.e., $n=2.04$. The critical value decided at the 0.005 level of significance is 0.0156, which is less noteworthy than the D-max value and henceforth, the watched authorship information distribution holds good for the Lotka's law, and consequently, the Lotka`s law for the open access PCOS literature research at global level acknowledge for the Authorship distributions.

SIGNIFICANT KEYWORDS :
The distribution of the keywords in the open access publication output of polycystic ovary syndrome at global level can be seen from Map-1.Total number of 6354 keywords has been identified.
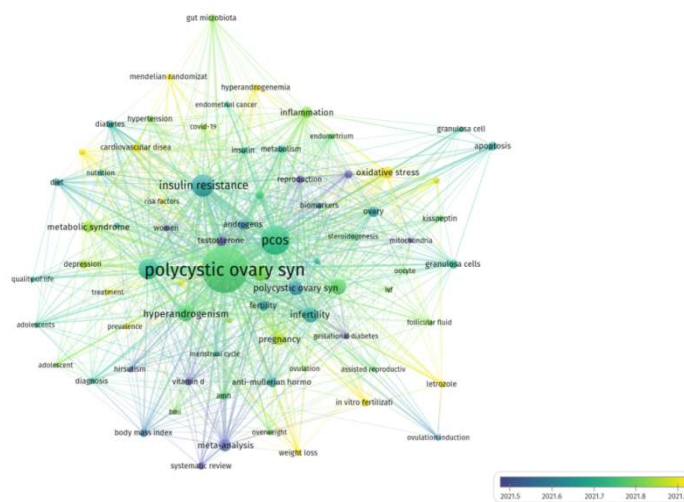
Source: Author`s own work   **Map-1: significant keywords**

The word 'women' has repeatedly been used 1280 times by PCOS scientists and it is dominated in the first rank. The word 'polycystic-ovary-syndrome' has been used 1136 times, which stood in the second rank in the words Occurrences. The word 'insulin-resistance' is occupied in the third rank with used constant frequently 826 times. The word 'prevalence' has 507 frequencies followed by "Risk" with 407 and "Obesity" frequency of 306. Moreover, the following nineteen words have used frequency at above hundred times.

CONCLUSION :

The maximum frequency of Open Access publications on PCOS is found in the form of Article representing publications (67%). 3065 (73.802%) publications pertaining to PCOS have been disseminated through the Green mode of open access .English is the dominant language with 99.80% publications. It is found that top ten institutions are occupied by China and are playing the leading role in publishing the Open Access scholarly content on PCOS. Kolmogorov Simonov test is applied for the wellness of the Lotka's law for the estimations of Lotka's types acquired from least square methods. The outcomes show that the estimation of D-max, i.e., 0.0156 decided with Lotka's type, i.e., n=2.04. The critical value decided at the 0.005 level of significance is 0.0156, which is less noteworthy than the D-max value and henceforth, the watched authorship information distribution holds good for the Lotka's law, and consequently, the Lotka`s law for the open access PCOS literature research at global level acknowledge for the Authorship distributions.

REFERENCES :

1.Goodarzi, M. O., Dumesic, D. A., Chazenbalk, G., & Azziz, R. (2011). Polycystic ovary syndrome: Etiology, pathogenesis and diagnosis. Nature Reviews Endocrinology, 7(4), 219-231.

2. Azziz, R., Carmina, E., Chen, Z., Dunaif, A., Laven, J. S., Legro, R. S., & Lizneva, D. (2016). Polycystic ovary syndrome. *Nature Reviews Disease Primers*, 2(1), 16057.

3.Ahmad, Muneer, Batcha, M. Sadik, and Jahina, S. Roslein (2019), "Testing Lotka's law and pattern of author productivity in the scholarly publication of Artificial Intelligence", *Library Philosophy and Practice* (e-journals).

4. Su, Y.-S., Lin, C.-L., Chen, S.-Y., & Lai, C.-F. (2020). Bibliometric study of social network analysis literature. *Library Hi Tech,* 38(2), 420-433. https://doi.org/10.1108/LHT-01-2019-0028

5. Barik, N., & Jena, P. (2021). Author productivity pattern and applicability of Lotka's inverse square law: A bibliometric appraisal of selected LIS open access journals. *Digital Library Perspectives, 37*(3), 223-241. https://doi.org/10.1108/DLP-10-2020-0103

6. Fallah, M., Fahimifar, S., Noruzi, A., & Ghorbi, A. (2023). Application of Lotka's law and i10-index with the number of authors of articles in chemistry in Iran published between 2000 and 2020. *Informology*, 2(1), 63-76.