

Digital epidemiology and geographic mapping of G6PD deficiency: retrospective analytic of trends database existing on the surface web



Ahmed Al-Imam^{1, 2, 3, 4}

¹Department of Anatomy and Cellular Biology, College of Medicine, University of Baghdad, Iraq, ²Department of Postgraduate Medicine, School of Life and Medical Sciences, University of Hertfordshire, United Kingdom, ³The Canadian Association for Neuroscience, ⁴The Japanese Association of Anatomists

Submitted: 02-07-2018

Revised: 29-07-2018

Published: 01-09-2018

ABSTRACT

Background: Glucose-6-phosphate dehydrogenase (G6PD) deficiency is an inherited X-linked recessive condition in which the body does not synthesise a sufficient quantity of the G6PD enzyme. Hemolytic anaemic episodes occur following exposure to some medications, foods, or even infections in G6PD-deficient individuals. **Aims and Objectives:** To assess the digital epidemiology as well as the geographic mapping of G6PD deficiency in the world including countries of the Mediterranean basin. **Materials and Methods:** This study is primarily based on a digital epidemiological analysis and geographic mapping based on data retrieved from Google Trends database, a *very large database*, existing on the surface web. A retrospective analysis was carried out for the period from the 22nd of March 2015 to the 24th of January 2016. Trends data will also be contrasted with a collateral data set of a cross-sectional analytic for the same period which was conducted based on the exploration of cases of G6PD deficiency admitted to Jordan University Hospital. **Results:** Concerning geographic mapping, countries of the Basin of the Mediterranean Sea and the Arabian Gulf accounted for one-third of the entire geographic map at 15.66% and 18.18% respectively. Countries of the Mediterranean basin included Jordan, Italy, Lebanon, Israel, Egypt, Greece, Syria, Cyprus, Tunisia, Algeria, and Morocco. Contributing countries surrounding the Arabian Gulf included Kuwait, Kingdom of Saudi Arabia, Bahrain, Iran, United Arab Emirates, Qatar, and Iraq. The Hashemite Kingdom of Jordan contributed to 2.78% of the global map which accounted for 17.74% of the entire Mediterranean basin. Concerning digital epidemiology, the prevalence was recorded as highest during May 2015 (12.10%), August 2015 (11.17%), and November 2015 (11.28%). Concerning the percentile contribution of monthly records, data were in harmony for those cases admitted to hospital and signals recorded via Google Trends. Both datasets averaged a percentile contribution of approximately 9% per month. **Conclusion:** This study is the first inferential research on G6PD deficiency based on data from a trends database and parallel data from Jordan University Hospital. Ambitious future research should deploy the use of machine learning for real-time and predictive analytics which will provide an excellent value for public health services and epidemiological studies.

Key words: Hemolytic anemia, Glucosephosphate dehydrogenase deficiency, Trends databases, Google trends, Retrospective studies

INTRODUCTION

Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency is the most common enzymatic disorder of genetic basis affecting the red blood cells (RBCs) of nearly 400 million individuals worldwide.^{1,2} It has a high prevalence in people of African, Asian, and Mediterranean descent.

Therefore, the severity of G6PD deficiency can vary significantly among ethnic groups.² The World Health Organization (WHO) has classified G6PD deficiency into variants (class-I to class-V) based on the level of enzyme deficiency and the severity of hemolytic episode and its clinical manifestations. Class-I deficiency represents the most severe form of G6PD deficiency. G6PD deficiency

Access this article online

Website:

<http://nepjol.info/index.php/AJMS>

DOI: 10.3126/ajms.v9i5.20495

E-ISSN: 2091-0576

P-ISSN: 2467-9100

Address for Correspondence:

Dr Ahmed Al-Imam, House 18/5, Al-Akhtal Street, District 318, Al-Adhamyia, 10053, Baghdad, IRAQ.

E-mail: tesla1452@gmail.com

© Copyright AJMS

in the region of the Mediterranean basin is usually a class-II deficiency, while class-III deficiency is due to defect in the gene G6PD A —. Class IV and Class V are of no clinical significance.³

The G6PD enzyme is part of the pentose monophosphate shunt. It catalyses the oxidation of glucose-6-phosphate and the reduction of nicotinamide adenine dinucleotide phosphate (NADP+) to nicotinamide adenine dinucleotide phosphate (NADPH). NADPH maintains glutathione in its reduced form, which acts as a scavenger for oxidative metabolites.³ G6PD deficiency is inherited as an X-linked recessive disorder like that of haemophilia and colourblindness.^{2,4} The condition is generally asymptomatic, but the atmospheric exposure or ingestion of fava beans and certain chemicals including drugs, as well as oxidant stress, bacterial or viral infectious agents can trigger an acute episode of hemolytic anaemia.⁴ Infections are the most common triggers of RBC hemolysis via a reactive inflammatory response leading to the generation of free radicals in macrophages, which can diffuse into the RBCs causing oxidative damage.^{4,5} Fortunately, symptoms of G6PD deficiency, including cardio-respiratory distress and pallor, typically disappear once the trigger is recognised or removed in parallel with management with blood transfusion and hemodynamic support for the patient.^{2,3,5} G6PD deficiency appears to be a risk factor for the development of diabetes mellitus.^{1,5}

This study will investigate, in an aim to reach a statistical inference, the digital epidemiology and the relevant geographic mapping of G6PD deficiency in the world, and to assess the contribution of the region of the Middle, the Arabian Gulf, and the Basin of the Mediterranean Sea.

MATERIALS AND METHODS

This study is purely observational, and it relies on a digital epidemiologic exploration based on retrospective analytics of Google Trends database from the 22nd of March 2015 to the 24th of January 2016.^{6,7} There will also be a collateral inclusion of some data from cross-sectional analytic of cases of G6PD deficiency admitted and diagnosed for the 1st time at Jordan University Hospital in Amman, the capital city of Jordan. All cases were admitted to the hospital within the specified timeframe. All individuals were of the Arab ethnicity from the Middle East who are permanent residents of Jordan. Data collection was based on a personal interview (5-10 minutes) conducted by a professional specialist pharmacist and a physician who has been well-trained to avoid biases including selection and interpretation errors.

The retrospective analytics was implemented via extraction of data already indexed within the very large database (VLDB) of Google Trends.⁸⁻¹⁰ These data are based on automated data gathering from millions of users existing on the surface web. The data was retrieved for the same period during which the cross-sectional analytics and interviews were conducted at Jordan University Hospital with an aim to compare regional data from Jordan and wide scale representative data from the entire world (surface web). Geographic mapping was also attempted via retrospective analysis aiming at assessing the fractional (percentile) contribution of Jordan, the Mediterranean Basin, and the Arabian Gulf region. Statistical analyses, descriptive and inferential, were carried out via *Microsoft Excel 2016* and the *Statistical Package for Social Sciences (SPSS v.20)*. The implemented statistical tests included *Student's t-test*, *the Analysis of Variance and Covariance (ANOVA)*, *Chi-Square test*, *regression models*, and *multivariate analysis*. An alpha value (α) of 0.05 and a confidence interval of 95% (95% CI) were considered as the cut-off margin for statistical inference.

The level-of-evidence for this study is estimated to be of level-3 in compliance with the categorisation scheme of level-of-evidence imposed by the *Oxford Center for Evidence-Based Medicine*.¹¹⁻¹³ A systematic review of the literature was carried out on medical and paramedical databases of literature including-PubMed, the Cochrane Library, Embase, EBSCO, Google Trends, Research Gate, Academia, Google Trends, and Sci-Hub. The grey literature databases were also consulted. The reviewed literature of interest was subsequently scrutinised for credibility and reliability via the implementation of the appropriate tools of critical appraisal.¹⁴

RESULTS

Based on data retrieved from Google Trends (Table 1), the highest digital epidemiology was recorded during May 2015 (12.10%), August 2015 (11.17%), and November 2015 (11.28%). Concerning the percentile contribution of monthly records, data were consistent, no statistically significant difference in monthly average, in between those cases admitted to Jordan University Hospital and data signals conveyed via Google Trends. Both datasets averaged a percentile contribution of approximately 9% per month. However, linear regression was inconclusive as it displayed a weak negative correlation without a statistical significance in between the two sets of data of hospital versus Google Trends (R^2 Score= 0.064) (Figure 1).

Geographic mapping via Google Trends was informative (Figures 1 and 2). The mapping included an impressive count of 69 countries. Ghana and Maldives were ranked

Table 1: Data on G6PD Deficiency from Jordan University Hospital versus Google Trends					
Jordan University Hospital			Google Trends (surface web)		
Time of Case Admission	Count of Cases	Percentile Contribution (%)	Count (interest)	Percentile Contribution (%)	
March, 2015	8	12.50	178	4.86	
April, 2015	3	4.69	342	9.34	
May, 2015	4	6.25	443	12.10	
June, 2015	9	14.06	326	8.90	
July, 2015	4	6.25	319	8.71	
August, 2015	0	0.00	409	11.17	
September, 2015	5	7.81	317	8.66	
October, 2015	12	18.75	320	8.74	
November, 2015	10	15.63	413	11.28	
December, 2015	4	6.25	297	8.11	
January, 2016	5	7.81	297	8.11	

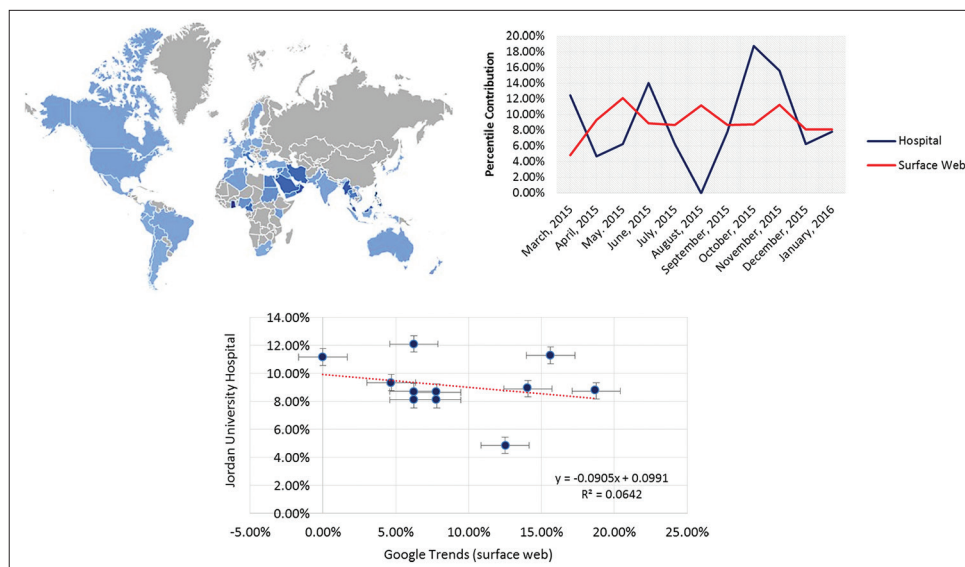


Figure 1: Jordan University Hospital versus Google Trends: Digital Epidemiology of G6PD Deficiency (top left), Percentile Contribution (top right), and Linear Regression (bottom)

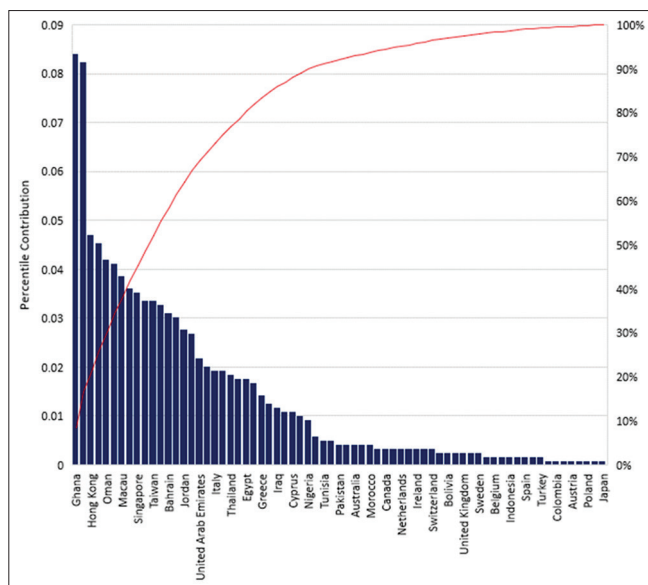


Figure 2: Geographic Mapping Based on Data from Google Trends

1st and 2nd (8.42% and 8.25%), while Peru, Colombia, Chile, Austria, South Africa, Poland, Argentina, and Japan all ranked last (0.08% for each country). Countries of the Mediterranean Basin and the Arabian Gulf Region accounted for one-third of the entire geographic map at 15.66% and 18.18% respectively. The top twenty-three contributing countries accounted for 80.39%. Those countries included Ghana, Maldives, Hong Kong, Philippines, Oman, Myanmar (Burma), Macau, Malaysia, Singapore, Kuwait, Taiwan, Saudi Arabia, Bahrain, Iran, Jordan, Cameroon, United Arab Emirates, Qatar, Italy, Lebanon, Thailand, Israel, and Egypt (Figure 2). Countries of the Mediterranean basin included Jordan, Italy, Lebanon, Israel, Egypt, Greece, Syria, Cyprus, Tunisia, Algeria, and Morocco. Contributing countries surrounding the Arabian Gulf included Kuwait, Kingdom of Saudi Arabia, Bahrain, Iran, United Arab Emirates, Qatar, and Iraq. The Hashemite Kingdom of Jordan contributed to

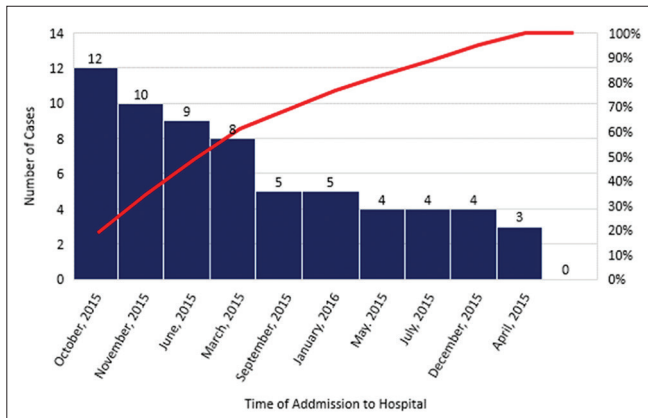


Figure 3: Descriptive Parameters: Number of Cases versus Timing of Interview (time of hospitalisation)

2.78% of the global map which accounted for 17.74% of the entire Mediterranean basin.

The total number of patients included in the cross-sectional analytic at Jordan University Hospital was sixty-four ($n=64$). All were from the Arab ethnicities inhabiting the capital city of Amman in the Hashemite Kingdom of Jordan. Male patients accounted for 62.5%. Approximately, eighty percent (80%) of patients were of an age ranging in between six months to five and a half years (5.5 years). The youngest patient was six months old while the oldest was eleven years old (Figure 3). Most patients (76.56%) were admitted to the hospital within 24 hours from onsets of symptomatic manifestations of the acute hemolytic anaemia, while the rest (21.88%) were admitted within 48 hours, and only one patient (1.56%) was admitted within 12 hours. Most of the patients were interviewed first at the ER (53.13%), department of paediatrics (42.19%), or at the paediatrics outpatient clinic (4.69%).

DISCUSSION

Future studies should attempt to venture with experimental and quasi-experimental study designs, including therapeutics and gene therapeutics in parallel with observational analyses. Thus, will actively promote the level-of-evidence of future research. Exploration of more than one trends database can be attempted as a representative of the surface web for rigour digital epidemiologic and geographic mapping. The incorporation of concepts of machine learning and pattern recognition for real-time and predictive analytics will provide an enormous impact of the level-of-evidence of studies as well as offer apreemptive warning system for mass casualties due to hemolytic anaemias clustered in a specific geographic location of the world.¹⁵ The benefits for public health and the economy will be colossal by sparing the work force, and other resources especially financial ones.

G6PD deficiency is not only the most common red blood cell enzymopathy but also represents the most common of all clinically significant enzyme defects in human biology.¹⁶ G6PD deficiency constitutes a significant public health problem in the Middle East.^{17,18} The highest incidence of G6PD deficiency has been described in Sardinia.¹⁹ G6PD deficiency is still a significant health problem in Jordan. The prevalence of G6PD deficiency in the north of Jordan was estimated to be around 5.5% while 17% of children admitted to princess Rahma teaching hospital with acute hemolytic anaemia due to G6PD deficiency were females.²⁰ G6PD deficiency is a public health problem that has been a predominant cause of severe hemolytic anaemia requiring urgent hospitalisation in young Palestinian children in Gaza strip.²¹ The highest prevalence of this enzymatic defect was reported among sub-Saharan African countries²². G6PD deficiency is also common in some provinces of Iran as fava beans represent an everyday food though Iran is considered as a moderately high incidence area for the G6PD deficiency.²³

CONCLUSION

Our study is the first of its kind to integrate a hybrid approach by extrapolating an inference based on data from trends databases in parallel with data from a major hospital in Jordan. The two sets of data were concordant and confirmatory for each other. Data from Google Trends displayed additional information on the geographic mapping and the fractional contribution of Jordan within the Mediterranean basin as well as the contribution of countries from the Arabian Gulf region. The study does have some limitations though, including reliance on pure observational analyses rather than an experimental approach. Data from Trends database may also be faulty due to the use of virtual private networks and internet protocol masking technologies which may affect the actual digital epidemiological assessment and the relevant geographic mapping. Additionally, the study did not incorporate the use of real-time or predictive analytics of data. The later can revolutionise epidemiological studies via the integration of concepts of big data in parallel with machine learning and automated data collection.

ACKNOWLEDGEMENTS

None.

REFERENCES

1. Walker BR and Colledge NR. Davidson's Principles and Practice of Medicine, 23rd ed. The Netherlands: Elsevier; 2018. Pages: 948 – 949.

2. Crook MA. *Clinical Biochemistry and Metabolic Medicine*, 8th edition. Hodder & Stoughton Ltd. 2012. Pages: 383
3. WHO Working Group (1989). "Glucose-6-phosphate dehydrogenase deficiency". *Bulletin of the World Health Organization*. 67 (6): 601–11.
4. Behrman REJ, Hal BK, Robert M, et al. *Nelson Essentials of Pediatrics*, 6th edition. Philadelphia. The Netherlands: Elsevier; 2011. Pages:- 570 – 571.
5. Millington DS, Sista R, Eckhardt A, Rouse J, Bali D, Goldberg R, et al. Digital microfluidics: a future technology in the newborn screening laboratory?. *In Seminars in perinatology* 2010; 34(2): 163-169.
6. Salathe M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. *PLoS computational biology* 2012;26;8(7):e1002616.
7. Bakker KM, Martinez-Bakker ME, Helm B and Stevenson TJ. Digital epidemiology reveals global childhood disease seasonality and the effects of immunization. *Proceedings of the National Academy of Sciences* 2016;113(24):6689-6694.
8. Choi H and Varian H. Predicting the present with Google Trends. *Economic Record* 2012;88(s1):2-9.
9. Carneiro HA and Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases* 2009;49(10):1557-1564.
10. Pelat C, Turbelin C, Bar-Hen A, Flahault A and Valleron AJ. More diseases tracked by using Google Trends. *Emerging infectious diseases* 2009;15(8):1327.
11. McAlister FA, Graham I, Karr GW and Laupacis A. Evidence-based medicine and the practicing clinician. *Journal of General Internal Medicine* 1999;14(4):236-242.
12. Abrams P, Khoury S and Grant A. Evidence-based medicine overview of the main steps for developing and grading guideline recommendations. *Progres en Urologie* 2007;17(3):681-684.
13. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) levels of evidence (background document). Oxford Center for Evidence-Based Medicine 2011; Sep 16.
14. Crowe M and Sheppard L. A review of critical appraisal tools show they lack rigor: alternative tool structure is proposed. *Journal of clinical epidemiology* 2011;64(1):79-89.
15. Nasrabadi NM. Pattern recognition and machine learning. *Journal of electronic imaging* 2007;16(4):049901.
16. Al-Sweedan SA, Jdaitawi H, Khriesat WM, Khader YY and Al-Rimawi HS. Predictors of severe hemolysis in patients with glucose-6-phosphate dehydrogenase deficiency following exposure to oxidant stresses. *Hematology/oncology and stem cell therapy* 2009;2(2):354-357.
17. Arese P, Gallo V, Pantaleo A and Turrini F. Life and death of glucose-6-phosphate dehydrogenase (G6PD) deficient erythrocytes—role of redox stress and band 3 modifications. *Transfusion Medicine and Hemotherapy* 2012;39(5):328-334.
18. Khaliq AA and Broyles RW. Hospital admissions: who is admitted through the emergency department?. *Health services management research* 2006;19(1):13-22.
19. Meloni T, Forteleoni G, Dore A and Cutillo S. Favism and hemolytic anemia in glucose-6-phosphate dehydrogenase-deficient subjects in North Sardinia. *Acta haematologica* 1983;70(2):83-90.
20. Al-Rimawi HS, Garaybah M, Al-Alimi AG and Amarin ZO. Glucose-6-Phosphate Dehydrogenase Activity in Jordanian Female Children with Acute Hemolytic Crises. *Jordan Medical Journal* 2007; 41(4).
21. Sirdah MM, Al-Kahlout MS and Reading NS. National G6PD neonatal screening program in Gaza Strip of Palestine: rationale, challenges and recommendations. *Clinical genetics* 2016;90(3):191-198.
22. Nkhoma ET, Poole C, Vannappagari V, Hall SA and Beutler E. The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis. *Blood Cells, Molecules, and Diseases* 2009;42(3):267-278.
23. Eghbalian F and Monsef AR. Evaluation of glucose-6-phosphate dehydrogenase deficiency without hemolysis in icteric newborns. *Iranian journal of pediatrics* 2007;17(1):36-40.

Author s Contribution:

AAI- Study design, Review of the literature, Statistical analysis and preparation of manuscript.

Work attributed to: The study was conducted in the College of Medicine at the University of Baghdad in Iraq.

Orcid ID:

Dr Ahmed Al-Imam -  <https://orcid.org/0000-0003-1846-9424>

Source of Funding: No external funding, this study was entirely self-funded. **Conflict of Interest:** None