Current Diabetes Reviews, 2018, *14*, 196-199

COMMENTARY

Diagnostic Validity and Clinical Utility of HbA1c Tests for Type 2 Diabetes Mellitus

Arthur H. Owora*

Public Health, Falk College, Syracuse University, New York, NY 13244, USA

ARTICLE HISTORY

Received: October 11, 2016
Revised: November 15, 2016
Accepted: November 24, 2016

DOI: 10.2174/1573399812666161129154559

1. INTRODUCTION

Individual and population level inference about risk and burden of diabetes is often made using diagnostic tests that are imperfect and prone to misclassification error (*i.e.* false positives and negatives) [1-3]. These errors or biases are rarely accounted for and could lead to inappropriate clinical decisions, inefficient allocation of scarce resources, and poor planning of disease prevention and treatment interventions [1]. The objective of this article is to describe how misclassification error due to imperfect diagnostic tests affects individual and population level inference, particularly involving the role of hemoglobin HbA1c (A1C) in diagnosing Type 2 Diabetes Mellitus (T2DM). An illustration of how disease prevalence, test sensitivity and specificity could be used by healthcare providers to inform individual level inference is also provided.

2. TYPE 2 DIABETES MELLITUS

T2DM is a prevalent metabolic disorder affecting about 9.3% the US population [4] and is projected to increase about 28% by the year 2050 if its incidence is not curtailed [5]. It is characterized by insulin resistance and/or low levels of insulin leading to abnormal glucose levels in the body. T2DM is associated with increased risk of retinopathy, nephropathy, neuropathy, and cardiovascular disease, peripheral arterial and cerebrovascular disease [6, 7]. In addition to causing substantial suffering and loss of work productivity to patients, these complications are also associated with significant caregiving burden. The annual healthcare costs associated with T2DM in the US are estimated to be $174 billion per year [8].

Given the burden of illness, there has been considerable public health interest in the early detection of T2DM risk to improve treatment prognosis. Despite the availability of effective early interventions (*e.g.* early pharmacotherapy, nutritional interventions augmented with physical activity and better lifestyle choices), [9] a substantial proportion of individuals with T2DM often receive late (or delayed) diagnoses [10]. One of the obstacles to early identification and diagnosis include inadequate healthcare provider knowledge of the diagnostic validity and potential clinical utility of existing point of care tests including the commonly-used A1c tests [11].

3. A1c TESTS AND DIAGNOSTIC CRITERIA FOR T2DM

The A1c test is the most common diagnostic and screening tool used for T2DM management and research [12, 13]. It measures an individual's average blood glucose level over the past three months; the higher an individual's blood glucose level, the higher the percentage test result. Unlike Fasting Plasma Glucose (FPG) tests which measure glucose floating free in the blood after fasting at the time of the test, the A1c test is less variable and reflects the average amount of glucose attached to hemoglobin over the past three months.

Recent T2DM diagnostic criteria proposed by the International Expert Committee based on the A1c suggest that levels ≥6.5% (48 mmol/mol) are indicative of T2DM and 6.0-6.4% identify those at high risk for progression toward T2DM [14]. The American Diabetes Association (ADA) also recognizes that A1c levels ≥6.5% indicate T2DM, while 5.7-6.4% indicate a high risk for progression to T2DM, even though these criteria are based upon both the FPG and 76-g Oral Glucose Tolerance Test (OGTT) [15]. Despite the numerous strengths that accompany the use of A1c (*e.g.*, highly-standardized with low intra-person variation, timely, fewer requirements for sample collection and storage), it exhibits poor validity and therefore remains a point of contention that limits its more widespread adoption (*e.g.* A1c test at 6.5% diagnostic threshold: Sensitivity=44%, Specificity=79%) [16-18]. Debate remains, with empirical research suggesting that a reliance strictly upon FPG and OGTT has also resulted in marked worldwide under-diagnoses and treatment of T2DM and pre-diabetes [19]. It has been long-recognized that the onset of T2DM often occurs years prior to both clinical diagnosis and treatment (often up to seven years), and that these delays are often associated with increased metabolic abnormalities, clinical manifestations, and

*Address correspondence to this author at the Public Health, Falk College, Syracuse University, 344E White Hall, 150 Crouse Dr. Syracuse, New York 13244 Public Health, United States; E-mail: ahowora@syr.edu

1875-6417/18 $58.00+.00

risk of death [20, 21]. The diagnostic validity of the A1c test has been examined at various diagnostic thresholds (cut-point levels) for the diagnosis of T2DM [22, 23]. Moreover, the choice of optimal diagnostic thresholds depended on whether the benefits of diagnosis and treatment outweighed harms for each specific patient population examined.

To illustrate, Buell, Kermah and Davidson (2010) conducted a study using the National Health and Nutrition Examination Survey (NHANES) (1999-2004: ≥20 years) and found that the sensitivity (*Se*) and specificity (*Sp*) of the A1c test (boronate affinity high performance liquid chromatography-HPLC) at a 5.8% level for the detection of T2DM ranged from 84% to 95% and 86% to 93% respectively [23]. At a 6.5% level, the *Se* and *Sp* were estimated to be 45% and 99% respectively [23]. Additionally, Kramer, Araneta, and Barret-Connor (2010) investigated older adults (mean age: 69; standard deviation: 11) estimated A1c *Se* and *Sp* at the 6.5% level to be 44% and 79% [22]. This study used a high-performance liquid chromatography A1c test with an automated analyzer [24]. These two studies suggest the *Se* and *Sp* estimates of A1c test could vary substantially not just based on the diagnostic threshold used but also the type of A1c test used.

Heterogeneity in the diagnostic performance of A1c tests has also been noted among patient groups with different ethnicities [25], ages [26], variants of hemoglobin – homoglobinopathies [27] and medical conditions such as HIV [28] and anemia [29] due to A1c-glucose discordance. This heterogeneity underscores the fact that the diagnostic performance of the A1c test is population specific and so is the optimal diagnostic threshold that is used to determine presence or absence of T2DM.

4. A1c TEST MISCLASSIFICATION ERROR

Based on the aforementioned, if such an A1c test was to be administered to 100 at risk individuals in a population with a '*true*' T2DM prevalence of 10% assuming a *Se* and *Sp* of 44% and 79%, one would expect 23 individuals to test positive (*i.e.* an apparent or biased prevalence of 23/100=23% *i.e*. [True positives (4) + False Positive (19)] =23), and 77 individuals to test negative. Therein, six patients with T2DM would be missed (false negative proportion of 66%: 1-*Se*) and 19 non-diabetic individuals would be wrongly classified as positive (false positive proportion of 21%: 1-*Sp*). With a positive predictive value of 17% (only 4 of 23 positive tests actually have T2DM), positive test results from A1c tests should be interpreted with caution to avoid the intangible burden and cost (*e.g.*, unnecessary worry and anxiety among positive tested individuals and the possibility of extra healthcare costs) due to further testing or initiation of unwarranted treatment/intervention. While the A1c test at the 6.5% diagnostic threshold may be good at ruling out diabetes (*i.e.* negative predictive value of 92% (71/77)) it wrongly classifies 21% (false positive proportion) of non-diabetic individuals as diabetic.

5. ACCOUNTING FOR A1c MISCLASSIFICATION ERROR USING THE LIKELIHOOD RATIO APPROACH

While the knowledge of test sensitivity and specificity is important, this alone is inadequate to inform a healthcare provider regarding the probability of 'true' disease status given an individual's test results. For sensitivity and specificity estimates to be useful to a healthcare provider, it is more theoretically-sound that estimates are combined with a likelihood ratio to better inform individual level inference (*i.e.*, probability of illness given a positive test result).

The likelihood ratio is a summary of how many times more likely (or less likely) an individual with disease is likely to have a particular test result than individuals without disease. By direct application of Bayes theorem, [30] likelihood ratios can be combined with results of A1c tests to estimate an individual's risk of disease (post-test probability). Therein, two variants of the likelihood ratios are needed, one for if an individual tests positive (positive likelihood ratio: LR+) and another if an individual tests negative (negative likelihood ratio: LR-). The LR+ is derived from dividing *Se* by 1-*Sp* (*i.e.*, *Se*/[1-*Sp*]) and the LR- from dividing 1- *Se* by *Sp* (*i.e.*, [1-*Se*]/*Sp*). The post-test probability can be derived from the post-test odds (*i.e.*, product of the pre-test odds and likelihood ratio as per Bayes theorem) as:

Post-test odds = pre-test odds x likelihood ratio (1);

and

Post-test probability = Post-test odds/ (1 + post-test odds) (2)

where:

odds = probability of being diabetic [p]/ (1-probability of being diabetic[p]) (3);

and

The probability of being diabetic[p] = odds/ [1 + odds] (4).

Based on this aforementioned example (*i.e.*, assuming *Se* is 44% and *Sp* is 79% and a prevalence of T2DM in a hypothetical population is 10%), the LR+ = 2.1 and LR- = 0.7 and pre-test odds = 0.1 (derived from 0.1/[1-0.1]), two scenarios may be developed to illustrate the use of the likelihood ratio.

Scenario one: If an individual tested positive for T2DM based on the A1c test, his/her post-test odds of being diabetic would be 0.1 X 2.1 =0.21 (*substitution into equation **1***: pre-test odds x LR+)

The post-test probability of being diabetic would be

= 0.21/ [1+0.21] - substitution into equation 4

= 0.17 (17%)

In interpreting these findings, these results indicate that after testing positive on the A1c, an individual's risk of having diabetes has increased by 7% (from 10% - the population prevalence). Further tests for T2DM might not be recommended among such patients unless more information regarding the individual's clinical history is known (*e.g.*, family history). Ideally, a higher post-test probability (depending on the threshold thought to be clinically relevant) would warrant further testing to confirm or rule out a diabetes diagnosis. However, recommendations for better dieting, increased physical activity and further monitoring are warranted as per the ADA guidelines [31].

Conversely, in addition to positive test result, if an individual had a family history indicative of higher T2DM risk.

His/her pre-test probability would be higher than 10% (population prevalence) since national prevalence estimates show that an individual with a family history of diabetes has a pre-test probability of diabetes of 14% (*i.e.* prevalence of diabetes among individuals with a family history of diabetes) [32]. Because such an individual's pre-test probability of diabetes is greater than 10% (*i.e.* the pre-test probability of an individual without a family history of diabetes), his/her post-test probability of being diabetic given a positive test result and a family history of diabetes is likely much higher than 17% (*i.e.* 23% derived by the direct application of multiplication theorem of probability for dependent events).

Scenario two: If an individual tested negative for diabetes based on the A1c test, his/her post-test odds of being diabetic would be 0.1 x 0.71 =0.07 (*substitution into equation 1*: pre-test odds x LR-)

The post-test probability of being diabetic would be

= 0.07/ [0.07 + 1] - substitution into equation 4

= 0.07 (7%)

According to these findings, and based upon existing ADA recommendations without any other indication for risk (*e.g.*, family history), such an individual may not necessarily warrant additional follow-up. It should be noted, however, that the above example is only for illustration purposes and is not a substitute for a full clinical workup and differential diagnoses, but hopefully augments that process for better clinical judgement among healthcare providers.

6. POPULATION LEVEL INFERENCE USING A1c TESTS

From a population-based research and policy perspective and often translating to individual case-level situations, studies that are reliant upon imperfect diagnostic tests (*i.e.*, those with markedly low sensitivity and specificity) to examine a disease risk or burden (prevalence) and associated risk factor measures of association (*e.g.*, odds ratios and risk ratios) should intuitively require adjustment for misclassification error to account for potential for false negatives and false positives. Applied specifically to T2DM, the valid estimation of an individual's T2DM risk is critical to appropriate clinical decision making. Unbiased prevalence estimates that are adjusted for misclassification error are better suited to inform appropriate allocation of scarce healthcare resources. Furthermore, unbiased measures of association between T2DM and suspected risk factors are critical to the design and implementation of preventive interventions.

CONCLUSION

The clinical relevance of population-specific diagnostic performance parameters such as sensitivity and specificity is not obvious; however, when combined with likelihood ratios they provide an intuitive and practical approach to making sense of an imperfect A1c test result for individual level inference. Current recommendations for improved inference from an individual's test result are predicated upon repeated blood glucose testing [33] including the use of estimated average glucose (eAG) [34]. In resource (*i.e.* time and testing materials) constrained primary care and community setting, such practices may not be cost-effective. Indeed, as illustrated above, with varying T2DM prevalence estimates in different populations, the interpretation of a test result could have different implications. Likelihood ratios can be used to take advantage of all available clinical information (*e.g.* family history) including knowledge of the heterogeneity of A1c test diagnostic performance across patient spectrums to arrive at more valid individual level inferences. The LR approach is especially useful when the A1c test is used as a screening tool in primary care and community settings where repeated testing for confirmatory diagnosis may not be feasible.

GUARANTOR STATEMENT

Owora, AH is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

AUTHOR CONTRIBUTIONS

A.O researched data, wrote manuscript, reviewed/edited the manuscript.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

A.O acknowledges Grant H. Skrepnek for comments provided on an earlier draft of this paper. A.O was supported by the Department of Public Health, Falk College, Syracuse University.

REFERENCES

[1] van den Bruel A, Cleemput I, Aertgeerts B, Ramaekers D, Buntinx F. The evaluation of diagnostic tests: Evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. J Clin Epidemiol 2007; 60(11): 1116-22.

[2] Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. J Clin Epidemiol 2009; 62(1): 5-12.

[3] Buysschaert M, Medina JL, Buysschaert B, Bergman M. Definitions (and Current Controversies) of Diabetes and Prediabetes. Curr Diabetes Rev 2016; 12(1): 8-13.

[4] Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014. Atlanta, GA: U.S.: Department of Health and Human Services, 2014.

[5] Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. Popul Health Metr 2010; 8: 29.

[6] Hanssen KF, Bangstad HJ, Brinchmann-Hansen O, Dahl-Jorgensen K. Blood glucose control and diabetic microvascular complications: Long-term effects of near-normoglycaemia. Diabet Med 1992; 9(8): 697-705.

[7] Fox CS, Coady S, Sorlie PD, *et al*. Increasing cardiovascular disease burden due to diabetes mellitus: the Framingham Heart Study. Circulation 2007; 115(12): 1544-50.

[8] American Diabetes Association. Economic costs of diabetes in the U.S. In 2007. Diabetes Care 2008; 31(3): 596-615.

[9] Greaves C, Sheppard K, Abraham C, *et al*. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011; 11(1): 119.

[10] Fraser L-A, Twombly J, Zhu M, *et al*. Delay in Diagnosis of Diabetes Is Not the Patient's Fault. Diabetes Care 2010; 33(1): e10.

[11] Bonora E, Tuomilehto J. The Pros and Cons of Diagnosing Diabetes With A1C. Diabetes Care 2011; 34(Supplement 2): S184-S90.

[12] Ang SH, Thevarajah M, Alias Y, Khor SM. Current aspects in hemoglobin A1c detection: A review. Clinica Chimica Acta 2015; 439: 202-11.

[13] Hirst JA, McLellan JH, Price CP, *et al*. Performance of point-of-care HbA1c test devices: implications for use in clinical practice - a systematic review and meta-analysis. Clinical chemistry and laboratory medicine. 2017; 55(2): 167-180.

[14] David MN. International expert committee report on the role of the a1c assay in the diagnosis of diabetes. Diabetes Care 2009; 32(7): 1327-34.

[15] American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 2010; 33(Suppl 1): S62-S9.

[16] Cowie CC, Rust KF, Byrd-Holt DD, *et al*. Prevalence of diabetes and high risk for diabetes using A1C criteria in the U.S. population in 1988-2006. Diabetes Care 2010; 33(3): 562-8.

[17] Kilpatrick ES, Bloomgarden ZT, Zimmet PZ. Is haemoglobin A1c a step forward for diagnosing diabetes? BMJ 2009; 339: b4432.

[18] Malkani S, Mordes JP. Implications of using hemoglobin A1C for diagnosing diabetes mellitus. Am J Med 2011; 124(5): 395-401.

[19] Alqahtani N, Khan WA, Alhumaidi MH, Ahmed YA. Use of glycated hemoglobin in the diagnosis of diabetes mellitus and prediabetes and role of fasting plasma glucose, oral glucose tolerance test. Int J Prev Med 2013; 4(9): 1025-9.

[20] Hillson RM, Hockaday TD, Newton DJ, Pim B. Delayed diagnosis of non-insulin-dependent diabetes is associated with greater metabolic and clinical abnormality. Diabet Med 1985; 2(5): 383-6.

[21] Samuels TA, Cohen D, Brancati FL, Coresh J, Kao WH. Delayed diagnosis of incident type 2 diabetes mellitus in the ARIC study. Am J Manag Care 2006; 12(12): 717-24.

[22] Kramer CK, Araneta MRG, Barrett-Connor E. A1C and diabetes diagnosis: The rancho bernardo study. Diabetes Care. 2010; 33(1): 101-3.

[23] Buell C, Kermah D, Davidson MB. Utility of A1C for Diabetes Screening in the 1999–2004 NHANES Population. Diabetes Care 2007; 30(9): 2233-5.

[24] Kramer CK, von Mühlen D, Gross JL, Laughlin GA, Barrett-Connor E. Blood pressure and fasting plasma glucose rather than metabolic syndrome predict coronary artery calcium progression: The rancho bernardo study. Diabetes Care 2009; 32(1): 141-6.

[25] Hare MJ, Magliano DJ, Zimmet PZ, *et al*. Glucose-independent ethnic differences in HbA1c in people without known diabetes. Diabetes Care 2013; 36(6): 1534-40.

[26] Pani LN, Korenda L, Meigs JB, *et al*. Effect of aging on A1C levels in individuals without diabetes: Evidence from the Framingham Offspring Study and the National Health and Nutrition Examination Survey 2001-2004. Diabetes Care 2008; 31(10): 1991-6.

[27] Rhea JM, Molinaro R. Pathology consultation on HbA(1c) methods and interferences. Am J Clin Pathol 2014; 141(1): 5-16.

[28] Kim PS, Woods C, Georgoff P, *et al*. A1C Underestimates Glycemia in HIV Infection. Diabetes Care 2009; 32(9): 1591-1593.

[29] Attard SM, Herring AH, Wang H, *et al*. Implications of iron deficiency/anemia on the classification of diabetes using HbA1c. Nutr Diabetes 2015; 5: e166.

[30] Efron B. Mathematics. Bayes' theorem in the 21st century. Science 2013; 340(6137): 1177-8.

[31] American Diabetes Association. *et al*. Standards of medical care in diabetes-2014. Diabetes Care. 2014; 37 (Suppl 1): S14-80.

[32] Annis AM, Caulder MS, Cook ML, Duquette D. Family history, diabetes, and other demographic and risk factors among participants of the National Health and Nutrition Examination Survey 1999-2002. Prev Chronic Dis 2005; 2(2): A19.

[33] American Diabetes Association. Classification and Diagnosis of Diabetes. Diabetes Care 2016; 39(Supplement 1): S13-S22.

[34] Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ. Translating the A1C assay into estimated average glucose values. Diabetes Care 2008; 31(8): 1473-8.