Table 2. Physician and Symptom Checkers' Diagnostic Accuracy, Stratified by the Acuity Level and Prevalence of the Conditions Described by the Clinical Vignettes

Vignette Characteristic	No. (%)		Listed, % (95% CI)			
			First[a]		Top 3[a]	
	Vignettes Completed by Human Dx Physicians	Vignettes Completed by Symptom Checkers	Human Dx Physicians	Symptom Checkers[b]	Human Dx Physicians	Symptom Checkers[a]
All vignettes	1105 (100)	770 (100)	72.1 (69.5-74.8)	34.0 (30.7-37.4)	84.3 (82.2-86.5)	51.2 (47.4-54.3)
Acuity level[c]						
High	398 (36.0)	263 (34.2)	79.1 (75.1-83.2)	24.3 (19.1-29.6)	89.2 (86.1-92.3)	39.5 (33.6-45.5)
Medium	376 (34.0)	260 (33.7)	70.7 (66.1-75.4)	37.7 (31.8-43.6)	84.3 (80.6-88.0)	56.9 (50.9-63.0)
Low	331 (30.0)	247 (32.1)	65.3 (60.1-70.4)	40.5 (34.3-46.7)	78.5 (74.1-83.0)	57.5 (51.3-63.5)
Vignette prevalence[d]						
Common	639 (57.8)	457 (59.4)	69.6 (66.1-73.2)	38.1 (33.6-42.5)	83.3 (80.4-86.2)	55.6 (51.6-60.7)
Uncommon	466 (42.2)	313 (40.6)	75.5 (71.6-79.5)	28.1 (23.1-33.1)	85.8 (82.7-89.0)	44.7 (38.4-49.3)

Totals may not add up to 100% owing to rounding.

[a] $P<.001$ for all comparisons between physicians and symptom checkers.

[b] Results described by Semigran et al.[4] Full version of clinical vignettes available at: http://www.bmj.com/content/bmj/suppl/2015/07/07/bmj.h3480.DC1/semh025489.ww1.pdf.

[c] Acuity level of vignettes defined by Semigran et al.[4] Differences across physicians and across symptom checkers for this category were statistically significant ($P < .001$).

[d] We defined "common" diagnoses as those that accounted for more than 0.3% of ambulatory visits (or >3 764 082 visits) in the United States in 2009 to 2010. These totals were compiled from data gathered by the Centers for Disease Control and Prevention, the National Ambulatory Medical Care Survey, and the National Hospital Ambulatory Medical Care Survey. Differences across physicians and across symptom checkers for this category were statistically significant ($P < .05$) except for the difference between the rate that physicians listed the correct diagnosis in the top 3 for common vs uncommon vignettes.

Across physicians, they were more likely to list the correct diagnosis first for high-acuity vignettes (vs low-acuity vignettes) and for uncommon vignettes (vs common vignettes). In contrast, symptom checkers were more likely to list the correct diagnosis first for low-acuity vignettes and common vignettes (Table 2).

Discussion | In what we believe to be the first direct comparison of diagnostic accuracy, physicians vastly outperformed computer algorithms in diagnostic accuracy (84.3% vs 51.2% correct diagnosis in the top 3 listed).[4] Despite physicians' superior performance, they provided the incorrect diagnosis in about 15% of cases, similar to prior estimates (10%-15%) for physician diagnostic error.[5] While in this project we compared diagnostic performance, future work should test whether computer algorithms can augment physician diagnostic accuracy.[6]

Key limitations included our use of clinical vignettes, which likely do not reflect the complexity of real-world patients and did not include physical examination or test results. Physicians who chose to use Human Dx may not be a representative sample of US physicians and therefore may differ in diagnostic accuracy. Symptom checkers are only 1 form of computer diagnostic tools, and other tools may have superior performance.

Hannah L. Semigran, BA
David M. Levine, MD, MA
Shantanu Nundy, MD
Ateev Mehrotra, MD, MPH

Author Affiliations: Harvard Medical School, Boston, Massachusetts (Semigran, Mehrotra); Brigham and Women's Hospital, Boston, Massachusetts (Levine); The Human Diagnosis Project, Washington, DC (Nundy).

Corresponding Author: Ateev Mehrotra, MD, MPH, Health Care Policy, Harvard Medical School, 180 Longwood Ave, Boston, MA 02115 (mehrotra@hcp.med.harvard.edu).

Published Online: October 10, 2016. doi:10.1001/jamainternmed.2016.6001

Author Contributions: Dr Mehrotra had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: All authors.
Acquisition, analysis, or interpretation of data: Semigran, Levine, Nundy.
Drafting of the manuscript: Semigran.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Semigran.
Administrative, technical, or material support: Levine, Mehrotra.
Study supervision: Nundy, Mehrotra.

Conflict of Interest Disclosures: Dr Nundy is an equity holder of The Human Diagnosis Project, the creators of Human Dx. No other disclosures are reported.

1. The National Academies of Science Engineering and Medicine. *Improving Diagnosis in Health Care*. Washington, DC: The National Academies Press; 2015.

2. Topol EJ. The Future of Medicine Is in Your Smartphone. *The Wall Street Journal*; January 9, 2015; The Saturday Essay.

3. Poon K, Okin PM, Kligfield P. Diagnostic performance of a computer-based ECG rhythm algorithm. *J Electrocardiol*. 2005;38(3):235-238.

4. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*. 2015;351:h3480.

5. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf*. 2013;22(suppl 2):ii21-ii27.

6. Bond WF, Schwartz LM, Weaver KR, Levick D, Giuliano M, Graber ML. Differential diagnosis generators: an evaluation of currently available computer programs. *J Gen Intern Med*. 2012;27(2):213-219.

LESS IS MORE

Prediabetes Risk in Adult Americans According to a Risk Test

The Diabetes Prevention Program and other studies found that individuals with impaired glucose tolerance (based on a 75-g oral glucose tolerance test) can decrease their risk of type 2

Copyright 2016 American Medical Association. All rights reserved.

Downloaded From: http://jamanetwork.com/ by a Tufts Univ. Hirsh Health Sciences Library User on 12/09/2016

diabetes developing either by an intensive supervised lifestyle intervention, including diet and exercise modification, or by metformin hydrochloride treatment.[1,2] Subsequently, the glycemic criteria for prediabetes were expanded to include hemoglobin A_{1c} and a decreased level for fasting glucose.[3] Although the benefit of type 2 diabetes prevention is unclear in this broader group, the Centers for Disease Control and Prevention, American Diabetes Association, and American Medical Association have promoted a web-based risk test to evaluate people at high risk for prediabetes for whom they recommend practice-based laboratory testing.[4] We estimated the proportion of the adult, nondiabetic US population that would be classified as being at high risk for prediabetes according to this widely endorsed risk instrument.

Editor's Note page 1863

Methods | Using data from the 2013-2014 National Health and Nutrition Examination Survey population older than 18 years without type 2 diabetes, we calculated risk scores for prediabetes based on 7 questions[4]:

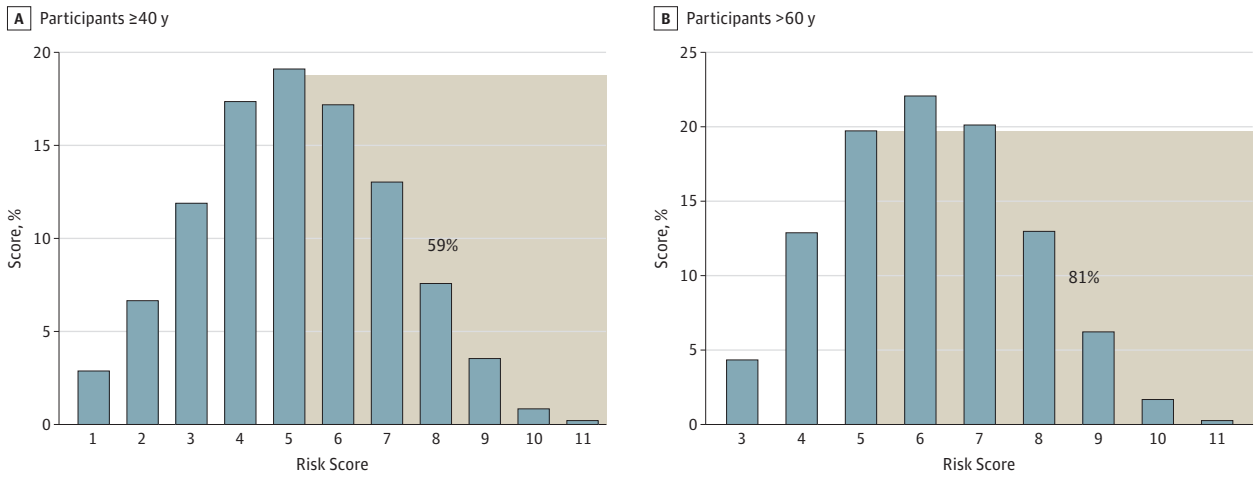
1. How old are you (1-3 points)?
2. Are you a man or a woman (1 point)?
3. If you are a woman, have you ever been diagnosed with gestational diabetes (1 point)?
4. Do you have a mother, father, sister, or brother with diabetes (1 point)?
5. Have you ever been diagnosed with high blood pressure (1 point)?
6. Are you physically active (1 point)?
7. What is your weight status (1-3 points)?

We selected the questions from the National Health and Nutrition Examination Survey that closely matched those in the instrument. Physical inactivity was defined by a negative answer to a set of 5 questions regarding activity level. We inferred results of the weighted proportions to the US adult population after a complete case analysis. The study was considered exempt from the institutional review board approval by Tufts Medical Center. The National Health and Nutrition Examination Survey obtained patient consent in a written form. Patients receiving a score of 5 or more are at high risk for prediabetes and are advised to visit their physician for a blood glucose test.

Results | Of 10 175 participants, 96.5% provided complete information for all questions. The **Figure** shows the distribution of the risk scores. Among people 40 years or older, the estimated number evaluated as being at high risk for prediabetes was 73.3 million, corresponding to 58.7% (sample size for the age group, 3815; 95% CI, 56%-62%) of the population (Figure, A). Among those participants older than 60 years, the weighted proportion of the population at high risk for prediabetes was 80.8% (sample size for the age group, 1841; 95% CI, 78%-84%) (Figure, B).

Discussion | When applied to the US population, the Centers for Disease Control and Prevention, American Diabetes Association, and American Medical Association risk instrument categorizes 3 of 5 people 40 years or older and 8 of 10 individuals 60 years or older as being at high risk for prediabetes, requiring a medical visit and a blood glucose test for confirmation. Given the expanded criteria, many of these high-risk individuals will have prediabetes when tested.[3] However, such a widespread process may be premature for many reasons. First, intensive lifestyle methods—even for those participants with impaired glucose tolerance—are most beneficial for those at the highest risk.[5] Second, according to the US Preventive Services Task Force, there is no direct evidence that type 2 diabetes prevention alters the risk for diabetes-related complications.[1,6] Third, to our knowledge, the natural history of prediabetes based on the latest American Diabetes Association criteria has not been prospectively assessed, but it

Figure. Risk Scores of Prediabetes Among the Nondiabetic Population in the United States According to Age



Distribution of scores for risk for prediabetes among the nondiabetic population in the United States 40 years or older (A) or older than 60 years (B). Shaded area identifies at-risk population.

Copyright 2016 American Medical Association. All rights reserved.

Downloaded From: http://jamanetwork.com/ by a Tufts Univ. Hirsh Health Sciences Library User on 12/09/2016

is likely that progression to type 2 diabetes will be slower with the expanded criteria compared with impaired glucose tolerance. Finally, medicalization of prediabetes may have the unintended consequence of reducing health care access to patients with type 2 diabetes and other chronic conditions. A valid method to examine for prediabetes should avoid unnecessary medicalization by labeling a disease predecessor as a medical condition and seek to concentrate on people at highest risk to allow for efficient distribution of limited health care resources.

Saeid Shahraz, MD, PhD
Anastassios G. Pittas, MD
David M. Kent, MD

Author Affiliations: Predictive Analytics and Comparative Effectiveness Center, Tufts Medical Center, Boston, Massachusetts (Shahraz, Kent); Division of Endocrinology, Tufts Medical Center, Boston, Massachusetts (Pittas).

Corresponding Author: Saeid Shahraz, MD, PhD, Predictive Analytics and Comparative Effectiveness Center, Tufts Medical Center, 800 Washington St, Campus Box 63, Boston, MA 02111 (shahraz@gmail.com).

Published Online: October 3, 2016. doi:10.1001/jamainternmed.2016.5919

Author Contributions: Drs Shahraz and Kent had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: All authors.
Acquisition, analysis, or interpretation of data: Shahraz, Kent.
Critical revision of the manuscript for important intellectual content: All authors.
Administrative, technical, or material support: Shahraz, Kent.
Study supervision: All authors.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by research grants DK076092 and DK098245 from the National Institute of Diabetes and Digestive and Kidney Diseases and the National Institutes of Health Office of Dietary Supplements (Dr Pittas).

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

1. Knowler WC, Barrett-Connor E, Fowler SE, et al; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002;346(6):393-403.

2. Balk EM, Earley A, Raman G, Avendano EA, Pittas AG, Remington PL. Combined diet and physical activity promotion programs to prevent type 2 diabetes among persons at increased risk: a systematic review for the Community Preventive Services Task Force. *Ann Intern Med*. 2015;163(6):437-451.

3. Yudkin JS, Montori VM. The epidemic of pre-diabetes: the medicine and the politics. *BMJ*. 2014;349:g4485.

4. Do you have prediabetes? Prediabetes Risk Test. https://doihaveprediabetes.org/pdf/Prediabetes_PrintableRiskTest(English).pdf. Accessed May 15, 2016.

5. Sussman JB, Kent DM, Nelson JP, Hayward RA. Improving diabetes prevention with benefit based tailored treatment: risk based reanalysis of Diabetes Prevention Program. *BMJ*. 2015;350:h454.

6. U.S. Preventive Services Task Force. Screening for type 2 diabetes mellitus in adults: U.S. Preventive Services Task Force recommendation statement [published correction appears in *Ann Intern Med*. 2008;149(2):147]. *Ann Intern Med*. 2008;148(11):846-854.

Editor's Note

The Medicalization of Common Conditions

When *JAMA Internal Medicine* launched the Less Is More series 6 years ago, we commented that one area of concern was "medicalization" of common conditions.[1] In this issue, Shahraz et al[2] elegantly demonstrate how common conditions can be "medicalized." Using NHANES data they find that a widely promoted web-based risk test would label more than 73 million Americans, including more than 80% of those older than 60 years, as being at high risk for "prediabetes," a condition never heard of 10 years ago.[2] We suggest a better approach to preventing the epidemic of obesity and its multiple health-related complications is emphasis on healthful diet, weight loss when appropriate, and increased physical activity at all levels—by schools, the medical profession, and public health and governmental agencies.

Rita F. Redberg, MD, MSc

1. Grady D, Redberg RF. Less is more: how less health care can result in better health. *Arch Intern Med*. 2010;170(9):749-750. doi:10.1001/archinternmed.2010.90

2. Shahraz S, Pittas AG, Kent DM. Prediabetes risk in adult Americans according to a risk test [published online October 3, 2016]. *JAMA Intern Med*. doi:10.1001/jamainternmed.2016.5919

Estimating 1-Year Mortality for High-Risk Primary Care Patients Using the "Surprise" Question

Palliative care improves the value of care for seriously ill patients, but resource constraints necessitate targeting palliative care interventions to patients who need them most.[1] The "surprise" question (SQ)—"Would you be surprised if this patient died in the next 12 months?"—has emerged as an attractive, simple solution for identifying patients who might benefit from palliative care.[2,3] Despite optimism about the potential of the SQ to identify primary care patients who would benefit from palliative care,[4] there is no evidence on its performance in this setting.

Methods | We identified patients screened for a high-risk care management program at a large academic primary care practice for whom the primary care physicians answered the SQ between August 30, 2012, and February 27, 2014. We assumed a no answer represented physician prediction of high 1-year mortality risk. Our primary outcome was mortality 1 year after SQ response, determined by linkage to Social Security Administration data. We obtained demographics and comorbidities[5] from electronic health records. We assessed SQ performance for estimating 1-year mortality using area under the receiver operating characteristic curve, sensitivity, positive predictive value, and odds ratio of a no response for 1-year mortality using univariate logistic regression. To quantify incremental benefit of the SQ for predicting 1-year mortality over and above routinely collected administrative data, we calculated the integrated discrimination improvement[6] of adding the SQ response to multivariate logistic regression of mortality on age, sex, and comorbidity score.[5] The institutional review board of Partners HealthCare approved this study with a waiver of informed consent. Analysis is based on patient data from medical records and Social Security Administration data (identified records); tabulations, test characteristics, and regression analyses are deidentified data.

Results | A total of 1737 patients were included in the study. Patients were predominantly female (1041 [60.3%]). Mean age

Copyright 2016 American Medical Association. All rights reserved.

Downloaded From: http://jamanetwork.com/ by a Tufts Univ. Hirsh Health Sciences Library User on 12/09/2016