

# FINDRISC in Latin America: a systematic review of diagnosis and prognosis models

Rodrigo M Carrillo-Larco <sup>1,2,3</sup>, Diego J Aparcana-Granda,<sup>2</sup> Jhonatan R Mejia,<sup>4</sup> Antonio Bernabé-Ortiz <sup>2,5</sup>

**To cite:** Carrillo-Larco RM, Aparcana-Granda DJ, Mejia JR, *et al.* FINDRISC in Latin America: a systematic review of diagnosis and prognosis models. *BMJ Open Diab Res Care* 2020;**8**:e001169. doi:10.1136/bmjdr-2019-001169

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2019-001169>).

DJA-G and JRM contributed equally.

Received 31 December 2019  
Revised 21 January 2020  
Accepted 22 February 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY. Published by BMJ.

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

<sup>2</sup>CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru

<sup>3</sup>Instituto de Investigación, Universidad Católica Los Ángeles de Chimbote, Chimbote, Peru

<sup>4</sup>Facultad de Medicina Humana, Universidad Nacional del Centro del Perú, Huancayo, Peru

<sup>5</sup>Universidad Científica del Sur, Lima, Peru

## Correspondence to

Dr Rodrigo M Carrillo-Larco; [rcarrill@ic.ac.uk](mailto:rcarrill@ic.ac.uk)

## ABSTRACT

This review aimed to assess whether the FINDRISC, a risk score for type 2 diabetes mellitus (T2DM), has been externally validated in Latin America and the Caribbean (LAC). We conducted a systematic review following the CHARMS (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies) framework. Reports were included if they validated or re-estimated the FINDRISC in population-based samples, health facilities or administrative data. Reports were excluded if they only studied patients or at-risk individuals. The search was conducted in Medline, Embase, Global Health, Scopus and LILACS. Risk of bias was assessed with the PROBAST (Prediction model Risk of Bias ASsessment Tool) tool. From 1582 titles and abstracts, 4 (n=7502) reports were included for qualitative summary. All reports were from South America; there were slightly more women, and the mean age ranged from 29.5 to 49.7 years. Undiagnosed T2DM prevalence ranged from 2.6% to 5.1%. None of the studies conducted an independent external validation of the FINDRISC; conversely, they used the same (or very similar) predictors to fit a new model. None of the studies reported calibration metrics. The area under the receiver operating curve was consistently above 65.0%. All studies had high risk of bias. There has not been any external validation of the FINDRISC model in LAC. Selected reports re-estimated the FINDRISC, although they have several methodological limitations. There is a need for big data to develop—or improve—T2DM diagnostic and prognostic models in LAC. This could benefit T2DM screening and early diagnosis.

## INTRODUCTION

With an increasing load in terms of prevalence,<sup>1</sup> disability and mortality,<sup>2–4</sup> as well as economic burden,<sup>5</sup> type 2 diabetes mellitus (T2DM) is a global threat to population health and health systems, especially in low-income and middle-income countries.<sup>1–5</sup> Although universal health coverage should secure treatment for all patients, this goal may not be realistic where there is not universal screening and where resources are limited to identify all at-risk populations. Therefore, inexpensive yet reliable screening tools could be useful to identify T2DM cases or high-risk people. Risk scores, both diagnostic and prognostic, help

identify people at high risk of having or developing T2DM. This way, these people could undergo further diagnostic tests, primary prevention or receive pharmacological treatment as needed. Nonetheless, risk scores need to be tested, and possibly adapted (ie, recalibrated), to produce accurate estimates to inform health decisions. Several T2DM risk scores have been developed,<sup>6–9</sup> although very few for Latin America and the Caribbean (LAC), where those available exhibit major limitations hindering their implementation across countries or their endorsement by policies or guidelines.<sup>10</sup> A well-known T2DM risk score is the FINDRISC,<sup>11</sup> which is also acknowledged by the Latin American diabetes guidelines as an available diabetes screening tool;<sup>12</sup> yet it is unknown if this model has been appropriately adapted in LAC. Consequently, we aimed to describe and assess if external validations of the FINDRISC model in LAC were conducted following adequate methods.<sup>13–14</sup> We will complement the available evidence about T2DM risk scores in LAC<sup>10</sup> and inform regional guidelines,<sup>12</sup> while also pinpointing research priorities and policies for T2DM screening and early diagnosis through risk stratification.<sup>15–16</sup>

## METHODS Protocol

This systematic review and critical appraisal of the scientific literature adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (online supplementary material). We followed the CHARMS (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies) methodology to formulate the review framework, research question and strategy (table 1).<sup>17–18</sup>

## Information sources

The search strategy was conducted in five search engines: Embase, Medline and Global

**Table 1** CHARMS criteria to define research question and strategy

Concept	Criteria
Prognostic or diagnostic?	The review focused on FINDRISC regardless if it was studied as a diagnostic or prognostic model for T2DM.
Scope	To inform physicians, researchers and the general population whether they are likely to have T2DM (ie, diagnostic) or will be likely to have T2DM (ie, prognostic). FINDRISC models could be used for research, screening and treatment allocation in primary prevention.
Type of prediction modeling studies	<ul style="list-style-type: none"> <li>▶ Diagnostic/prognostic models with external validation.</li> <li>▶ Diagnostic/prognostic models without external validation.</li> <li>▶ Diagnostic/prognostic model validation.</li> </ul>
Target population to whom the prediction model applies	General adult population in LAC.
Outcome to be predicted	T2DM.
Time span of prediction	Prognostic models will not be included/excluded based on prediction time; that is, it could be short term (eg, next 2.5 years) or long term (eg, next 10 years).
Intended moment of using the model	FINDRISC models to be used in asymptomatic adults of LAC to assess their probability to have T2DM (ie, diagnostic) or their probability to develop T2DM in a predefined period (ie, prognostic).

Based on the CHARMS checklist.<sup>18</sup>

CHARMS, CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies; LAC, Latin America and the Caribbean; T2DM, type 2 diabetes mellitus.

Health through OVID, and also in Scopus and LILACS. The search was conducted on September 28, 2019. The search terms are available in the online supplementary material.

### Eligibility criteria

We sought FINDRISC models aiming to inform about the current (diagnostic) or future (prognostic) risk of T2DM in LAC populations. Selected original reports could have developed a new model using the same (or very similar) predictors as in the original FINDRISC<sup>11</sup>; similarly, they could have performed an independent external validation in LAC populations. The outcome of the diagnostic or prognostic FINDRISC models was T2DM. The outcome should have been ascertained with at least one biomarker (eg, fasting glucose, hemoglobin A1c (HbA1c) or oral glucose tolerance test). Thus, we did not include studies where the outcome relied entirely on self-reported diagnosis. We focused on adult men and women.

### Study selection

Original scientific reports were excluded if the study population only included people with a disease (eg, patients with hypertension) or based on a risk factor (eg, smokers). Similarly, studies with LAC populations in countries outside LAC were excluded (eg, Hispanics in the USA). Conversely, reports were included if they followed a probabilistic population-based sampling approach, were based on primary care settings, or were based on health or claims registries or administrative data. The original work should have focused on the FINDRISC model, regardless of whether they developed an identical new model, a very similar model, or independently externally validated the FINDRISC model. Studies were

included if they followed a cross-sectional or prospective observational design.

### Data collation process

We used EndNote and Rayyan<sup>19</sup> to remove duplicates from the search. First, we used Rayyan<sup>19</sup> to screen titles and abstracts, which were screened by two reviewers independently (pairwise combinations between RMC-L and DJA-G or JRM); discrepancies were solved by consensus. Then, two reviewers independently (pairwise combinations between RMC-L and DJA-G or JRM) studied the full text of those reports selected in the screening phase; discrepancies were solved by consensus. If consensus could not be reached, discrepancies were solved by a third party (AB-O).

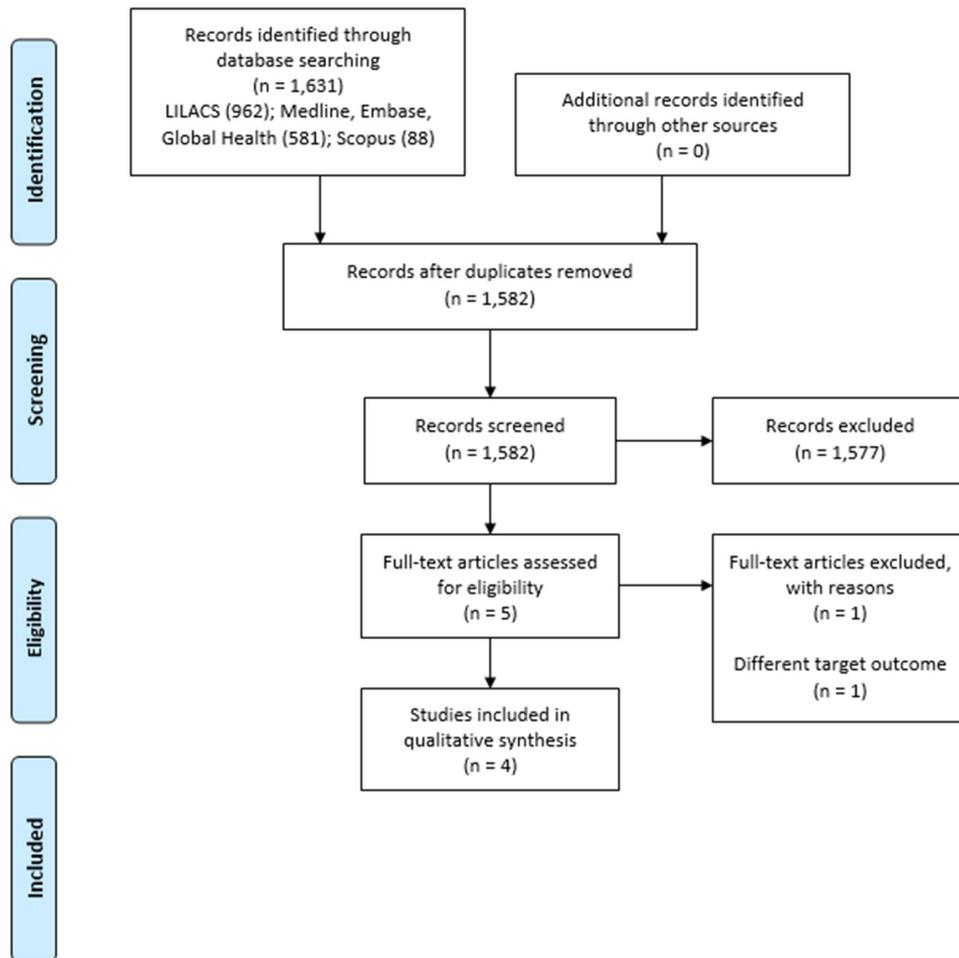
The authors developed a data extraction form based on international guidelines for systematic reviews of prognosis models<sup>17,18</sup> and on a previous systematic review on the subject.<sup>10</sup> The data extraction form was not modified during data collation. Information was extracted as presented in the original reports by two reviewers independently (pairwise combinations between RMC-L and DJA-G or JRM); discrepancies were solved by consensus.

### Risk of bias of individual studies

Using the PROBAST (Prediction model Risk of Bias ASsessment Tool) tool for risk of bias assessment of prognosis models,<sup>20,21</sup> two reviewers (DJA-G and JRM) independently assessed the risk of bias of the selected reports. If there were any discrepancies, these were solved by consensus or by a third party (RMC-L).

### Synthesis of results

Because of the limited numbers of results and the great heterogeneity among them, only a qualitative synthesis was conducted.



**Figure 1** Flow chart of the study selection process.

## RESULTS

### Reports selection

The screening process included 1582 titles and abstracts, of which 1577 were excluded. Therefore, five reports were studied in full text. One report was excluded because they assessed a different outcome.<sup>22</sup> Finally, four reports (n=7502) were included in the qualitative review (figure 1).<sup>23–26</sup>

### General characteristics

All the reports were from South America: one from Peru,<sup>24</sup> one from Venezuela,<sup>26</sup> and two from Colombia.<sup>23, 25</sup> Bernabe-Ortiz *et al*<sup>24</sup> as well as Nieto-Martínez *et al*<sup>26</sup> studied population-based samples while Gomez-Arbelaez *et al*<sup>25</sup> and Barengo *et al*<sup>23</sup> analyzed data from health centers (online supplementary material).

The largest sample size was studied by Nieto-Martínez *et al* (n=3061); this was also a national representative sample.<sup>26</sup> The smallest sample was studied by Gomez-Arbelaez and colleagues (n=772).<sup>25</sup> The studied samples tended to include slightly more women than men, except for one report (58.3% men).<sup>25</sup> The mean age ranged from 29.5 to 49.7 years (table 2, online supplementary material).<sup>24, 25</sup>

Across reports, T2DM was ascertained with a combination of fasting plasma glucose  $\geq 126$  mg/dL, HbA1c  $\geq 6.5\%$  or 2-hour plasma glucose  $\geq 200$  mg/dL (table 2, online

supplementary material).<sup>23–26</sup> Undiagnosed T2DM prevalence was largest in the report by Barengo *et al* (5.10%; n=105),<sup>23</sup> followed by the study in Peru (4.70%; n=71),<sup>24</sup> the report from Venezuela (3.30%; n=101)<sup>26</sup> and the work by Gomez-Arbelaez *et al*<sup>25</sup> (2.59%; n=20) (table 2, online supplementary material).

### Predictors and modeling

None of the studies conducted an independent external validation of the FINDRISC model. Conversely, they used the same (or very similar) predictors to fit a new model.<sup>23–26</sup> In so doing, they all produced new coefficients and baseline risks.<sup>23–26</sup> As in the original FINDRISC,<sup>11</sup> numeric variables were categorized. The modeling strategy was consistently logistic regression, and complete-case analyses were conducted (online supplementary material).

Two authors developed new risk scores.<sup>23, 24</sup> Barengo *et al*<sup>23</sup> started with nine candidate predictors to develop a Colombian version of the FINDRISC with six predictors; predictor selection was based on univariate analysis (online supplementary material). Bernabe-Ortiz *et al*<sup>24</sup> developed a simplified version of the FINDRISC including 5 predictors, yet there were 12 candidate predictors selected through stepwise backward elimination (online supplementary material).

Table 2 General characteristics

Study	Outcome prevalence (%)	Mean age (years)	Men (%)	Outcome details	Baseline sample size	Outcome events (n)	Outcome events per candidate predictors
Gomez-Arbelaiz <i>et al</i> <sup>25</sup>	2.59	58.34	29.53	The diagnosis of type 2 diabetes mellitus was established when fasting plasma glucose $\geq 126$ mg/dL, OGTT $\geq 200$ mg/dL and/or HbA1c $\geq 6.5\%$ .	772	20	
Bernabe-Ortiz <i>et al</i> <sup>24</sup> FINDRISC Simplified, FINDRISC and LA- FINDRISC	4.70	48.20	49.70	Individuals who were not aware of having type 2 diabetes mellitus and had fasting glucose $\geq 126$ mg/dL ( $\geq 7.0$ mmol/L) or 2-hour plasma glucose $\geq 200$ mg/dL ( $\geq 11.1$ mmol/L).	1609	71	5.92
Nieto-Martinez <i>et al</i> <sup>26</sup> LA- FINDRISC and FINDRISC	3.30	39.90	46.97	Diabetes was defined if the fasting plasma glucose was $\geq 126$ mg/dL or if the 2-hour OGTT glucose was $\geq 200$ mg/dL.	3061	101	
Barengo <i>et al</i> <sup>23</sup> ColDRISC and FINDRISC	5.10	47.20	38.00	Individuals who had fasting plasma glucose $\geq 126$ mg/dL or 2-hour plasma glucose $\geq 200$ mg/dL were classified as having T2DM.	2060	105	11.67

HbA1c, hemoglobin A1c; OGTT, oral glucose tolerance test.

### Model performance

None of the studies reported any calibration metrics (online supplementary material).<sup>23–26</sup> Conversely, they all focused on discrimination (area under the receiver operating curve) and other classification metrics, including sensitivity, specificity, and positive and negative predictive values. The area under the receiver operating curve was consistently above 65.0% (table 3, online supplementary material).

### Risk of bias

All studies exhibited high risk of bias mainly due to limitations in the analytical approach, for example limited number of outcome events. In this line, a complete-case analysis was consistently preferred versus multiple imputation. Most importantly, calibration metrics were consistently not reported. On the other hand, participants, predictors, and outcome criteria showed low risk of bias. There was low applicability concern (table 4, online supplementary material).

## DISCUSSION

### Main findings

This review did not find any independent external validations of the original FINDRISC model in LAC, as the four reports herein described re-estimated the FINDRISC model; in other words, authors computed new coefficients and baseline risks instead of using the original ones to test the model performance in a new population with subsequent recalibration if needed. While the analyzed reports exhibited methodological limitations, including reduced number of outcome events and not reporting calibration metrics, they showed acceptable discrimination performance. In LAC, risk prediction research needs to be improved to generate reliable tools for risk stratification, which could offer a cost-effective approach in the cascade to identify new and future T2DM cases.<sup>27</sup>

### Limitations of the review

Although we followed a comprehensive methodology, there are still limitations to be acknowledged. First, we did not search gray literature; however, we would not expect results from these sources, if any, to substantially change the main findings or conclusions of this review. Second, the focus of this work was on LAC; whether our findings apply to other world regions mostly hosting low-income and middle-income countries deserves further verification.

### Limitations of the selected reports

We have previously pinpointed several methodological limitations of T2DM risk scores in LAC,<sup>10</sup> and these would also apply to those herein studied. Although there is literature addressing good methods for development and validation of risk scores,<sup>13 14</sup> the most recurrent pitfall herein identified is the limited number of outcome events, which may allow including few predictors or could lead to overfitting of the prediction model. We understand

**Table 3** Performance metrics

First author, and assessed model	Discrimination (%)	Classification measures
Gomez-Arbelaez <i>et al</i> <sup>25</sup>	74.77 (95% CI 57.22 to 92.32) (men) and 71.75 (95% CI 58.68 to 84.81) (women)	At a cut-off of $\geq 14$ . Men: sensitivity=66.7; specificity=75.2; positive predictive value=6.8; negative predictive value=98.8; Youden's index=0.419. Women: sensitivity=71.4; specificity=62.6; positive predictive value=4.8; negative predictive value=98.8; Youden's index=0.340.
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> FINDRISC Simplified	71.10	At a cut-off of 3: sensitivity=0.859; specificity=0.467; positive predictive value=0.074; negative predictive value=0.985; likelihood ratio positive=1.6; likelihood ratio negative=0.3; diagnostic OR=5.3.
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> FINDRISC	69.00	At a cut-off of 11: sensitivity=0.690; specificity=0.668; positive predictive value=0.094; negative predictive value=0.978; likelihood ratio positive=2.1; likelihood ratio negative=0.5; diagnostic OR=4.5.
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> LA-FINDRISC	68.00	At a cut-off of 10: sensitivity=0.704; specificity=0.591; positive predictive value=0.079; negative predictive value=0.970; likelihood ratio positive=1.7; likelihood ratio negative=0.5; diagnostic OR=3.4.
Nieto-Martinez <i>et al</i> , <sup>26</sup> LA-FINDRISC	72.2 (95% CI 66.8 to 77.5) (men) and 72.40 (95% CI 63.9 to 81.0) (women)	For men at a cut-off of 9: sensitivity=72.2; specificity=62.2; positive likelihood ratio=1.91. For women at a cut-off of 10: sensitivity=71.4; specificity=65.4; positive likelihood ratio=2.06.
Nieto-Martinez <i>et al</i> , <sup>26</sup> FINDRISC	72.90 (95% CI 67.6 to 78.1) (men) and 73.20 (95% CI 64.8 to 81.6) (women)	
Barengo <i>et al</i> , <sup>23</sup> CoIDRISC	74 (95% CI 70 to 79)	At a cut-off of 4: sensitivity=0.73; specificity=0.67; positive predictive value=0.106; negative predictive value=0.979.
Barengo <i>et al</i> , <sup>23</sup> modified FINDRISC	73 (95% CI 69 to 78)	At a cut-off of 10: sensitivity=0.72; specificity=0.60; positive predictive value=0.084; negative predictive value=0.984.

that (big) data with enough outcome events may be scarce in LAC; thus, we value and acknowledge the available research. Recently, methods have been developed to define sample size for risk prediction models with binary outcomes.<sup>28</sup> Where possible, researchers could adhere to these standards.

We strongly believe there is a great need to look for (big) data, for example, national surveys (eg, WHO STEP-wise approach to Surveillance (STEPS) or Demographic and Health Surveys (DHS)). These surveys are available

in many countries, and pooling them, following adequate techniques, could generate a rich database to develop a T2DM risk score for LAC. Finally, it is also worrying that none of the studies reported calibration metrics such as calibration slope, calibration in the large or calibration plots.<sup>13 14</sup> Calibration refers to the agreement between observed and predicted events.<sup>13 14</sup> Therefore, it provides information to understand whether the model is underestimating (observed > predicted) or overestimating (observed < predicted) the outcome. As risk prediction

**Table 4** Risk of bias assessment of individual diagnostic/prediction models (PROBAST)<sup>20 21</sup>

First author and assessed model	RoB				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Barengo <i>et al</i> , <sup>23</sup> CoIDRISC	+	+	+	-	+	+	+	-	+
Barengo <i>et al</i> , <sup>23</sup> modified FINDRISC	+	+	+	-	+	+	+	-	+
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> FINDRISC Simplified	+	+	+	-	+	+	+	-	+
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> FINDRISC	+	+	+	-	+	+	+	-	+
Bernabe-Ortiz <i>et al</i> , <sup>24</sup> LA-FINDRISC	+	+	+	-	+	+	+	-	+
Gomez-Arbelaez <i>et al</i> <sup>25</sup>	+	+	+	-	+	+	+	-	+
Nieto-Martinez <i>et al</i> , <sup>26</sup> LA-FINDRISC	+	+	+	-	+	+	+	-	+
Nieto-Martinez <i>et al</i> , <sup>26</sup> FINDRISC	+	+	+	-	+	+	+	-	+

+ indicates low RoB/low concern regarding applicability; - indicates high RoB/high concern regarding applicability. PROBAST, Prediction model Risk of Bias ASsessment Tool; RoB, risk of bias.

research further penetrates in LAC, standard and sound methods should be adopted and reported appropriately; thereby, robust tools will be available to be incorporated in health policies and guidelines.<sup>13 14 29</sup>

The summarized reports also provided metrics usually available for diagnostic tools, including positive/negative predictive values and positive/negative likelihood ratios. Of these additional metrics, the negative predictive value was the largest, consistently above 97%. This refers that a subject with a negative test is in fact disease-free. In other words, of 100 people who take the test and have a negative result, over 97 of them would not have the disease. This metric depends on the prevalence in the underlying population; that is, this is not an intrinsic property of the model. Thus, this metric would not be useful in generalizing the accuracy of the test across populations with different prevalences. Nonetheless, this could suggest that for people with a very low score or a score below a established threshold, further tests are not needed because they are most likely not to have diabetes.

### Clinical and public health relevance

The American<sup>30</sup> and Canadian<sup>31</sup> T2DM guidelines include risk scores to identify people who would need further laboratory tests to confirm T2DM;<sup>30 31</sup> these documents suggest specific risk scores such as the Canadian Diabetes Risk Assessment Questionnaire.<sup>31</sup> The LAC guidelines, on the other hand, support the use of risk scores for screening purposes, without advocating for any tool in particular, although they acknowledge the FINDRISC as a relevant and useful tool.<sup>12</sup> Probably the LAC guidelines do not support a risk score in particular due to the dearth of tools and the limitations of the few available ones.<sup>10</sup> Following the example of the US and Canadian guidelines, LAC T2DM institutions should support and foster the development of a strong T2DM risk score, which could benefit from national survey data or large pooling data endeavors.

Whether risk scores are the best method to screen for diabetes is yet to be known. Other alternatives include massive screening with blood tests (eg, random glucose or HbA1c) or screening based on single risk factors (eg, people with severe obesity). The first alternative may not be feasible in low-income and middle-income countries or rural settings, where costs and scarce laboratory facilities may preclude this option. Screening on single risk factors may not be sensible enough, hence the need for risk scores to combine several predictors to compute a more comprehensive probability. Also, there is evidence suggesting that screening with a risk stratification tool, such as a risk score, is a cost-effective approach.<sup>27</sup> While other screening methods are being developed and proven better than risk scores, risks scores need to be improved to provide accurate results that can inform public health (eg, number of people at risk in need of tests) and clinical medicine (eg, when to start counseling or treatment). The diabetes guidelines for Latin America do not explicitly recommend using the FINDRISC, yet

they signal the FINDRISC as a relevant screening tool.<sup>12</sup> In Colombia, on the other hand, clinical guidelines do recommend the FINDRISC.<sup>32</sup> Assessing which guidelines recommend the FINDRISC, or other risk scores, is beyond the scope of this work. However, given the limitations herein pinpointed as well as by a previous systematic review on the subject,<sup>10</sup> we recommend cautious use of available tools, particularly if they are being used in populations different from those used in developing the model.

### CONCLUSIONS

There has not been an external validation of the FINDRISC model in LAC, where several re-estimations of this model have been conducted. The available research has benefitted from studies with limited coverage, for example, small cross-sectional studies. This calls to strengthen the use of (big) data or national surveys across LAC to develop—or improve—T2DM diagnostic and prognostic risk scores. This could have large positive impact on T2DM screening and early diagnosis in LAC. Overall, the discrimination accuracy of the FINDRISC in LAC seems adequate, although no evidence is available on calibration metrics.

**Contributors** RMC-L and AB-O conceived the idea. RMC-L, DJA-G and JRM conducted the search and data extraction. DJA-G and JRM conducted the risk of bias. RMC-L wrote the manuscript with input from all coauthors. All authors approved the submitted version.

**Funding** The study received funding from Strategic Award, Wellcome Trust-Imperial College Centre for Global Health Research (100693/Z/12/Z), and Imperial College London Wellcome Trust Institutional Strategic Support Fund (Global Health Clinical Research Training Fellowship) (294834/Z/16/Z ISSF ICL). RMC-L is supported by a Wellcome Trust International Training Fellowship (214185/Z/18/Z). The funder had no role in the conception or conduct of this work, neither in the preparation of the results nor in manuscript writing. The authors alone are responsible for the results and opinions in this work.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** No human subjects were studied; thus, this review was classified as of low risk.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data sharing not applicable as no data sets generated and/or analyzed for this study. This is a systematic review of the scientific literature.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

### ORCID iDs

Rodrigo M Carrillo-Larco <http://orcid.org/0000-0002-2090-1856>  
Antonio Bernabé-Ortiz <http://orcid.org/0000-0002-6834-1376>

### REFERENCES

- 1 NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet* 2016;387:1513–30.
- 2 GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195

- countries and territories, 1980-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1736-88.
- 3 GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1859-922.
  - 4 GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1789-858.
  - 5 Bommer C, Sagalova V, Heeseemann E, et al. Global economic burden of diabetes in adults: projections from 2015 to 2030. *Diabetes Care* 2018;41:963-70.
  - 6 Abbasi A, Peelen LM, Corpeleijn E, et al. Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. *BMJ* 2012;345:e5900.
  - 7 Collins GS, Mallett S, Omar O, et al. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med* 2011;9:103.
  - 8 Juarez LD, Gonzalez JS, Agne AA, et al. Diabetes risk scores for Hispanics living in the United States: a systematic review. *Diabetes Res Clin Pract* 2018;142:120-9.
  - 9 Noble D, Mathur R, Dent T, et al. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163.
  - 10 Carrillo-Larco RM, Aparcana-Granda DJ, Mejia JR, et al. Risk scores for type 2 diabetes mellitus in Latin America: a systematic review of population-based studies. *Diabet Med* 2019;36:1573-84.
  - 11 Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003;26:725-31.
  - 12 Diabetes ALde. Guías ALAD sobre el Diagnóstico, Control y Tratamiento de la Diabetes Mellitus Tipo 2 con Medicina Basada en Evidencia - Edición, 2019URL. Available: [http://www.revistaalad.com/guias/5600AX191\\_guias\\_alad\\_2019.pdf](http://www.revistaalad.com/guias/5600AX191_guias_alad_2019.pdf)
  - 13 Riley RD, Dvd W, Croft P, et al. *Prognosis research in health care: concepts, methods, and impact Oxford*. UK: Oxford University Press, 2019.
  - 14 Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating: Springer international publishing*, 2019.
  - 15 World Health Organization. Noncommunicable diseases and mental health. ncd global monitoring framework. Available: [https://www.who.int/nmh/global\\_monitoring\\_framework/en/](https://www.who.int/nmh/global_monitoring_framework/en/)
  - 16 Pan American Health Organization / World Health Organization. Health policies: regulatory measures for diabetes prevention and control. Available: [https://www.paho.org/hq/index.php?option=com\\_docman&view=list&slug=health-policies-regulatory-measures-for-diabetes-prevention-control-8623&Itemid=270&lang=en](https://www.paho.org/hq/index.php?option=com_docman&view=list&slug=health-policies-regulatory-measures-for-diabetes-prevention-control-8623&Itemid=270&lang=en)
  - 17 Debray TPA, Damen JAAG, Snell KIE, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;356:i6460.
  - 18 Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.
  - 19 Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile APP for systematic reviews. *Syst Rev* 2016;5:210.
  - 20 Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1-33.
  - 21 Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51-8.
  - 22 Muñoz-González MC, Lima-Martínez MM, Nava A, et al. FINDRISC modified for Latin America as a screening tool for persons with impaired glucose metabolism in Ciudad Bolívar, Venezuela. *Med Princ Pract* 2019;28:324-32.
  - 23 Barengo NC, Tamayo DC, Tono T, et al. A Colombian diabetes risk score for detecting undiagnosed diabetes and impaired glucose regulation. *Prim Care Diabetes* 2017;11:86-93.
  - 24 Bernabe-Ortiz A, Perel P, Miranda JJ, et al. Diagnostic accuracy of the Finnish diabetes risk score (FINDRISC) for undiagnosed T2DM in Peruvian population. *Prim Care Diabetes* 2018;12:517-25.
  - 25 Gomez-Arbelaez D, Alvarado-Jurado L, Ayala-Castillo M, et al. Evaluation of the Finnish diabetes risk score to predict type 2 diabetes mellitus in a Colombian population: a longitudinal observational study. *World J Diabetes* 2015;6:1337-44.
  - 26 Nieto-Martínez R, González-Rivas JP, Ugel E, et al. External validation of the Finnish diabetes risk score in Venezuela using a national sample: the EVESCAM. *Prim Care Diabetes* 2019;13:574-82.
  - 27 Khunti K, Gillies CL, Taub NA, et al. A comparison of cost per case detected of screening strategies for type 2 diabetes and impaired glucose regulation: modelling study. *Diabetes Res Clin Pract* 2012;97:505-13.
  - 28 Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med* 2019;38:1276-96.
  - 29 Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Diabet Med* 2015;32:146-54.
  - 30 American Diabetes Association. 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes-2019*. *Diabetes Care* 2019;42:S13-28.
  - 31 Diabetes Canada Clinical Practice Guidelines Expert Committee. Diabetes Canada 2018 clinical practice guidelines for the prevention and management of diabetes in Canada. *Can J Diabetes* 2018;42:S1-325.
  - 32 Sistema General de Seguridad Social en Salud - Colombia. Guia de practica clinica para El diagnostico, tratamiento Y seguimiento de la diabetes mellitus tipo 2 en La poblacion mayor de 18 años. Available: [http://gpc.minsalud.gov.co/gpc\\_sites/Repositorio/Conv\\_637/GPC\\_diabetes/DIABETES\\_TIPO\\_2\\_COMPLETA.pdf](http://gpc.minsalud.gov.co/gpc_sites/Repositorio/Conv_637/GPC_diabetes/DIABETES_TIPO_2_COMPLETA.pdf)