BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Hepatitis C risk score as a tool to identify infected individuals: A demonstration study in Egypt

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-085506
Article Type:	Original research
Date Submitted by the Author:	20-Feb-2024
Complete List of Authors:	El-Khoury, Rayane; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group Chemaitelly, Hiam; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group; Weill Cornell Medicine - Qatar, World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis Alaama, Ahmed S.; World Health Organisation Regional Office for the Eastern Mediterranean, Department of Communicable Diseases Hermez, Joumana G.; World Health Organisation Regional Office for the Eastern Mediterranean, Department of Communicable Diseases Nagelkerke, Nico; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group Abu-Raddad, Laith; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar; Weill Cornell Medicine, Department of Population Health Sciences
Keywords:	EPIDEMIOLOGY, Risk Factors, Public health < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

ORIGINAL RESEARCH

Hepatitis C risk score as a tool to identify infected individuals: A demonstration study in Egypt

Rayane El-Khoury[1,2], Hiam Chemaitelly[1,2,3], Ahmed S. Alaama[4], Joumana G. Hermez[4], and Nico Nagelkerke[1], Laith J. Abu-Raddad[1,2,3,5,6*]

[1]*Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar, Cornell University, Qatar Foundation - Education City, Doha, Qatar*
[2]*World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis, Weill Cornell Medicine–Qatar, Cornell University, Qatar Foundation – Education City, Doha, Qatar*
[3]*Department of Population Health Sciences, Weill Cornell Medicine, Cornell University, New York, New York, USA*
[4]*Department of Communicable Diseases, HIV/Hepatitis/STIs Unit, World Health Organization Regional Office for the Eastern Mediterranean, Cairo, Egypt*
[5]*Department of Public Health, College of Health Sciences, Member of QU Health, Qatar University, Doha, Qatar*
[6]*College of Health and Life Sciences, Hamad bin Khalifa University, Doha, Qatar*

Word count: Abstract: 276 words, Main text: 3,380 words.
Number of tables: 3.
Number of figures: 3.
Running head: Hepatitis C virus risk score in Egypt.
Funding: The Qatar National Research Fund (NPRP 12S-0216-190094).

***Correspondence to:** Professor Laith J. Abu-Raddad, Infectious Disease Epidemiology Group, World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis, Weill Cornell Medicine - Qatar, Qatar Foundation - Education City, P.O. Box 24144, Doha, Qatar. Telephone: +(974) 4492-8321. Fax: +(974) 4492-8333. E-mail: lja2002@qatar-med.cornell.edu.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ABSTRACT

Objectives: Hepatitis C virus (HCV) infection poses a global health challenge. By the end of 2021, the World Health Organization estimated that less than a quarter of global HCV infections had been diagnosed. There is a need for a public health tool that can facilitate the identification of infected individuals and link them to testing and treatment, and that can be customized for each country.

Methods: We derived and validated a risk score to identify infected individuals in Egypt and demonstrated its utility. Utilizing data from the 2008 and 2014 Egypt Demographic and Health Surveys, two risk scores were constructed through multivariable logistic regression analysis. A range of diagnostic metrics was then calculated to evaluate the performance of these scores.

Results: The 2008 and 2014 risk scores exhibited similar dependencies on sex, age, and type of place of residence. Both risk scores demonstrated high and similar areas under the curve of 0.77 (95% CI: 0.76-0.78) and 0.78 (95% CI: 0.77-0.80), respectively. For the 2008 Risk Score, sensitivity was 73.7% (95% CI: 71.5-75.9%), specificity was 68.5% (95% CI: 67.5-69.4%), positive predictive value (PPV) was 27.8% (95% CI: 26.4-29.2%), and negative predictive value (NPV) was 94.1% (95% CI: 93.5-94.6%). For the 2014 Risk Score, sensitivity was 64.0% (95% CI: 61.5-66.6%), specificity was 78.2% (95% CI: 77.5-78.9%), PPV was 22.2% (95% CI: 20.9-23.5%), and NPV was 95.7% (95% CI: 95.4-96.1%). Each score was validated by applying it to a different survey database than the one used to derive it.

Conclusions: Implementation of HCV risk scores is an effective strategy to identify carriers of HCV infection and to link them to testing and treatment at low cost to national programs.

Keywords: Hepatitis C virus, viral hepatitis, risk score, Egypt, Demographic Health Survey.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

What is already known on this topic

The World Health Organization has set a global target to eliminate HCV infection as a public health problem by 2030. The Middle East and North Africa (MENA) region is the most affected by HCV infection. While mass testing and treatment programs may be relevant in countries with high prevalence, other countries exhibit relatively low HCV prevalence, rendering such programs less cost-effective. There is a need for a public health tool that can aid in identifying potentially infected individuals to link them to testing and treatment.

What this study adds

This study demonstrated the effectiveness of a risk score as a non-invasive public health tool comprising a few simple questions to identify carriers of HCV infection and link them to testing and treatment. Specifically demonstrated in Egypt, the tool exhibited good diagnostic accuracy, as indicated by various diagnostic performance metrics.

How this study might affect research, practice or policy

A single national survey for HCV infection can be sufficient to develop an effective risk score for HCV infection, which can become an integral component of the national strategy to eliminate this infection in a given country.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

INTRODUCTION

Hepatitis C virus (HCV) infection is a global public health challenge[1, 2] and a major cause of morbidity and mortality, resulting in liver cancer, fibrosis, and cirrhosis[3]. By end of 2021, the World Health Organization (WHO) estimated that 58 million people were infected with HCV, but only 15 million of them were diagnosed and only 9 million received treatment[4]. Direct-acting antivirals (DAA) offer highly effective treatment to cure this infection and to prevent progression toward severe forms of liver disease[5], as well as an opportunity to reduce HCV transmission through treatment as prevention[6, 7]. Accordingly, the WHO has set a global target to eliminate HCV infection as a public health problem by 2030[2, 8].

While DAAs are becoming accessible globally, it has been challenging to identify carriers of this infection so as to treat them, especially in the Middle East and North Africa (MENA), the region most affected by HCV infection and where most infected persons remain undiagnosed[9, 10]. Limited resources have made it challenging for viral hepatitis programs to find low-cost and cost-effective approaches to identify infected persons. While mass testing and treatment programs may be relevant in high prevalence countries, other countries have relatively low HCV prevalence making such programs less cost-effective[10-16]. While low-cost point-of-care tests (POCs) have been beneficial in some countries, such as Egypt[17], they remain relatively expensive for countries like Pakistan, which bear a substantial share of the global burden[18-20]. There is a need for a public health tool that can assist in identifying potentially infected persons so as to link them to testing and treatment.

One such tool is the use of risk scores to identify potentially infected individuals. A risk score comprises a small set of simple questions that can be used to assess the likelihood that an individual has a specific health condition[21-24], in this case, HCV infection. Such risk scores

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

have proven influential as public health tools for a range of health conditions, such as diabetes[21-24].

In this study, we demonstrate the application of this public health tool for HCV infection in Egypt, aiming to illustrate the public health value and practical utility of developing HCV risk scores in various countries. The risk score derived here is not intended for universal application across diverse settings; it is specifically designed for Egypt. However, the concept and analytical approach can be adapted to other countries by considering the local HCV epidemiology to determine the relevant factors and their respective weights for inclusion in a score tailored to each specific context.

METHODS

Egypt Demographic and Health Surveys

The Egypt Demographic and Health Survey (EDHS) is a national survey that collected data pertaining to the health and demographics of a nationally representative sample of the resident population of Egypt, including HCV infection[25, 26]. The EDHS that included HCV biomarkers was conducted in 2008 and 2014 and used rigorous sampling methods[27]. Details on study design, data collection, and laboratory methods can be found in El-Zanaty et al.[25, 26].

HCV antibody testing was done using a third generation enzyme-linked immunosorbent assay (ELISA), the Enzyme Immunoassay Adlatis EIAgen HCV Ab test (Adaltis Inc., Montreal, Canada)[25, 26]. All samples that were positive in the ELISA assay and 5% of the negative samples were then retested using a more specific assay, the chemiluminescent microplate immunoassay (CMIA ARCHITECT plus i1000SR, Abbott Diagnostic, USA)[25, 26]. If a sample was positive in both the ELISA and the CMIA testing, it was also tested for current

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

active infection, using real-time, reverse-transcription polymerase chain reaction (RT-qPCR) testing to detect HCV ribonucleic acid (RNA)[25, 26]. Samples were further retested for internal and external quality assurance[25, 26]. Here we restrict our analyses to the HCV antibody results.

Data from the EDHS 2008 and EDHS 2014 were downloaded with permission from Measure DHS[28]. The data can be accessed through application to the DHS Program at https://dhsprogram.com. For purposes of this study, the EDHS individual database was merged with the HCV biomarker database, based on established guidelines for managing DHS data[27]. All individuals with results for HCV antibody testing were included in the analysis.

Risk score derivation

Associations of HCV antibody positivity (seropositivity) with a priori variables that are easy to evaluate in a primary healthcare setting, and that can be included in a risk score, were investigated. These variables included sex (male versus female), age (5-year age strata), and type of place of residence (urban versus rural). Frequency distributions were generated to describe demographic and clinical profiles of tested individuals.

Chi-square tests and univariable logistic regression were implemented to investigate associations. Participants younger than 15 years of age were excluded as this age group was not included in the EDHS 2008 and has low HCV prevalence (Table S1)[6, 29-31]. Odds ratios (ORs), 95% confidence intervals (CIs), and p-values were reported. Covariates with p-values ≤0.1 in univariable regression analysis were considered possibly associated with HCV seropositivity. These were included in the multivariable analysis for estimation of adjusted ORs (AORs) and associated 95% CIs and p-values. No other forward or backward elimination for variable selection was used. Covariates with p-values ≤0.05 in the multivariable model were

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

considered predictors of HCV seropositivity. Univariable and multivariable analyses were adjusted for sampling weights.

A risk score was constructed based on the β-coefficients obtained from the multivariable regression model. β-coefficients were multiplied by a factor of 10 and then rounded to the nearest integer. The total risk score was calculated by adding the individual scores. To keep the score simple enough for use in primary healthcare and other general population settings, we did not consider any interaction terms.

Performance and validation of the risk score

A receiver operating characteristics (ROC) curve was plotted to investigate the performance of the risk score in predicting HCV seropositivity at different score cut-offs. A larger area under the curve (AUC), also called the c-index, indicates better performance of the risk score. The cut-off for the score was determined by maximizing the sum of the sensitivity and specificity. Sensitivity is the probability that the risk score will yield a positive diagnosis in a subject who is truly HCV antibody-positive. Specificity is the probability that the risk score will yield a negative diagnosis in a subject who is truly HCV antibody-negative.

Performance of the risk score was also investigated by estimating the positive predictive value (PPV) and the negative predictive value (NPV) of the risk score. PPV is the probability that a subject with a positive diagnosis per the risk score is truly HCV antibody-positive. NPV is the probability that a subject with a negative diagnosis per the risk score is truly HCV antibody-negative. The proportion of subjects who have scores greater than or equal to the cut-off of the risk score was estimated to determine the proportion of individuals that need to be biochemically tested for HCV antibodies.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

To validate the performance of the EDHS 2008 risk score, it was applied to the EDHS 2014 data, providing an independent validation with a dataset different from the one used for its derivation. Performance diagnostics were subsequently assessed. Given the pronounced cohort effect in the epidemiology of HCV infection in Egypt[6, 29-31], the age variable was adjusted to reflect the 6-year interval between the surveys. For example, individuals who were 11 years old in 2008 would have been 17 years old at the time of the second survey in 2014. The same approach was also used to validate the EDHS 2014 risk score—it was applied to the EDHS 2008 database and performance diagnostics were assessed.

While the cut-off for the score was determined by maximizing the sum of sensitivity and specificity, this cut-off can be adjusted as needed from a programmatic standpoint to optimize a specific diagnostic metric, such as sensitivity instead of specificity. To illustrate this flexibility, an additional analysis was incorporated featuring a variety of score cut-offs, resulting in diverse values of sensitivity, specificity, PPV, and NPV. Such additional analysis enables program managers and readers to discern the trade-offs among these diagnostic metrics and observe the implications of selecting an alternative programmatic approach, such as prioritizing the optimization of sensitivity over specificity.

Analyses were conducted in Stata version 16.1 (Stata Corporation, College Station, TX, USA). The study was reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Table S2).

RESULTS

In the 2008 EDHS, 11,126 individuals 15-59 years of age were tested, of whom 1,571 were antibody-positive[25]. The 2014 EDHS included children 1-14 years of age in addition to adults 15-59 years of age[26]. In this latter survey, 26,047 individuals were tested of whom 1,456 were

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

antibody-positive[26].

Characteristics of individuals who were tested for HCV antibodies and the proportion of each population stratum that was HCV antibody-positive are shown in Table S1 for both of the EDHS surveys. Results of both surveys were consistent, taking into account the age shift in the national cohort with the passage of 6 years between the EDHS 2008 and EDHS 2014.

HCV seropositivity was strongly associated with sex, age, and place of residence in both national surveys (Table 1 and Table S3). Male sex and rural residence were associated with higher seropositivity. Seropositivity increased rapidly with age.

The 2008 and 2014 Egypt Hepatitis C Risk Scores derived using the EDHS 2008 and EDHS 2014 data, respectively, are shown in Figure 1. The 2008 Risk Score had a range of 0-41. The 2014 Risk Score had a range of 0-53. Both showed similar dependence on sex, age, and type of place of residence. Both demonstrated high and similar AUCs (Figure 2). The AUC was 0.77 (95% CI: 0.76-0.78) for the 2008 Risk Score and 0.78 (95% CI: 0.77-0.80) for the 2014 Risk Score. The highest sum of sensitivity and specificity was obtained at a score cut-off value of 22 for the 2008 Risk Score and at a cut-off of 34.5 for the 2014 Risk Score.

For the 2008 Risk Score, sensitivity was 73.7% (95% CI: 71.5-75.9%), specificity was 68.5% (95% CI: 67.5-69.4%), PPV was 27.8% (95% CI: 26.4-29.2%), and NPV was 94.1% (95% CI: 93.5-94.6%) (Table 2). For the 2014 Risk Score, sensitivity was 64.0% (95% CI: 61.5-66.6%), specificity was 78.2% (95% CI: 77.5-78.9%), PPV was 22.2% (95% CI: 20.9-23.5%), and NPV was 95.7% (95% CI: 95.4-96.1%). The proportion of the population 15-59 years of age that needed to be biochemically tested for HCV antibodies was 37.2% (95% CI: 36.3-38.1%) using the 2008 Risk Score and 25.5% (95% CI: 24.9-26.2%) using the 2014 Risk Score. Of all HCV-infected persons in the EDHS samples, application of this score would have diagnosed (that is

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

identified; sensitivity) 73.7% (95% CI: 71.5-75.9%) and 64.0% (95% CI: 61.5-66.6%) of all HCV antibody-positive persons in samples of the EDHS 2008 and 2014, respectively.

When the 2008 Risk Score was applied to the EDHS 2014 data, the AUC was 0.75 (95% CI: 0.74-0.77), the sensitivity was 66.1% (95% CI: 63.5-68.6%), and the specificity was 72.3% (95% CI: 71.5-73.1%) (Table 2). These performance indicators were similar to the original performance indicators generated using the EDHS 2008 data, as well as to the performance indicators of the 2014 Risk Score on the EDHS 2014 data. Therefore, this application validates this risk score. A similar outcome was found when the 2014 Risk Score was applied to the EDHS 2008 data, also providing a validation of the 2014 risk score.

Figure 3 displays the proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections, using each of EDHS 2008 and EDHS 2014 data. The figure shows the effect of prioritization of testing for those with higher to lower risk score. This provides a demonstration of the utility of using the risk score: a large proportion of HCV infections can be diagnosed by testing only a small proportion of the population. It is most efficient programmatically to start testing individuals with the highest risk score and progressively moving on to those with lower and lower risk scores. As testing is expanded to those with low risk scores, the yield in identifying more HCV infections is very limited.

Table 3 illustrates the implications of selecting various score cut-offs, providing insight into the trade-offs among different diagnostic metrics, as well as the proportion of the population requiring biochemical testing and the proportion of all HCV-infected individuals identified through the application of this score. For instance, by enhancing the specificity of the risk score, the PPV increases, and the proportion of the population necessitating testing decreases. This

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

reduction in testing requirements helps alleviate costs and streamline the logistics of the test-and-treat program. However, this enhanced program efficiency comes at the expense of lower NPV and sensitivity, implying a smaller proportion of HCV-infected individuals in the population being identified through the risk score.

DISCUSSION

We demonstrated that a risk score that consists of few simple questions that are easy to evaluate in a primary healthcare setting or implemented through a website or an app that helps persons identify their risk of being HCV infected, provides an effective and non-invasive public health tool to identify carriers of HCV infection and to link them to testing and treatment. Biochemical testing methods to identify HCV infected persons are invasive and time-consuming and require human and financial resources, as well as complex logistics, making them less scalable, particularly in resource-limited settings. In contrast, initial screening using a risk score can be easily administered or self-administered, is non-invasive, and requires minimal resources and logistics. Therefore, HCV risk scores can be an indispensable strategy for the global response to attain the target of HCV elimination as a public health problem by 2030.

While the concept of a risk score shares similarities with risk-based testing, which has been implemented in some countries, predominantly in higher-income nations[32-34], the risk score approach transcends mere risk-based testing. It enables a broader application across various settings and situations and can significantly contribute to raising awareness of the infection among the general population. The risk score approach represents a tool that addresses several public health needs simultaneously, extending the application of risk-based testing beyond conventional healthcare settings. Moreover, it entails minimal costs and logistics, making it feasible even in resource-limited settings.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Remarkably, the risk score, comprising three simple questions, demonstrated considerable diagnostic accuracy, as evidenced by the values of various diagnostic metrics, including AUC, sensitivity, specificity, PPV, and NPV. Of particular note is the high NPV, ensuring that a negative result is highly unlikely to be a false negative, thereby obviating the need for individuals with a negative outcome using the score to undergo testing for HCV antibodies. The score also identified 73.7% and 64.0% of all HCV infections in the EDHS 2008 and EDHS 2014 samples, respectively. Thus, the score fulfills its objective of facilitating the efficient identification of individuals with HCV infection while minimizing the necessity for conducting biochemical testing. This underscores the value of this approach in identifying as many HCV-infected persons as early as possible and initiating treatment before progression to serious clinical disease.

This approach was demonstrated for Egypt, considering the availability of two EDHS surveys to derive and validate the score. The two scores exhibited comparable structures and diagnostic performances, with minor differences attributed to sampling variation of the same population across two distinct rounds of the EDHS surveys. Each score was validated by applying it to a database other than the one used to derive it. The latter application yielded a diagnostic performance that was comparable to the original diagnostic performance against the database used to originate it. This highlights how a single national survey for HCV infection may be sufficient to develop an effective risk score for this infection, and that can become an integral component of the national response to eliminate HCV infection.

The approach demonstrated in this study can be applied in other countries, including those in the MENA region. In countries where nationally representative population-based surveys have been conducted, these surveys can serve as the basis for deriving the risk score, as was done in this

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

study. However, only three MENA countries—Egypt, Libya, and Pakistan—have conducted such surveys[25, 26, 35, 36]. For countries where such surveys are not available[10-16], the risk score can still be derived using data from available regional surveys. Alternatively, if regional surveys are not available, the effects of risk factors for infection can be pooled, either in terms of odds ratios or relative risks, using data from analytical studies[37]. These effects can also be derived from meta-regression analyses applied to all available HCV prevalence studies for each country[38-45].

While this study focused on demonstrating the utility of this concept as a public health tool, actual application of this approach to different countries can be enhanced for even higher diagnostic accuracy. One extension could be adding more variables to the score in a manner tailored to the local epidemiology of each country. For instance, province or city of birth and/or current residence, prior exposure to an HCV mode of transmission[37], or history of HCV infection in the family, could be added, among others. Given that the risk of exposure to HCV infection varies immensely by at-risk population type and shows a distinctive hierarchy[46], an additional component to the score could be to integrate the at-risk population type as a variable[41, 46], thereby further enhancing the diagnostic accuracy of the score. Testing strategies, therefore, could be highly efficient in identifying HCV infected persons at a modest cost.

However, caution must be exercised to prevent the creation of stigma associated with HCV infection or the use of an HCV risk score. For instance, it may not be feasible to include questions about stigmatized behaviors in the MENA context, such as injecting drug use or specific sexual practices, when the score is applied in general population settings like primary healthcare. However, such questions may be appropriate in other settings, such as voluntary

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

counseling and testing (VCT) centers or outreach efforts by community organizations working with the most at-risk populations[47]. It is important also for the risk score to factor community acceptance in its design and implementation, ensuring it addresses the specific needs of certain groups, such as women of childbearing age in contexts where the risk of HCV vertical transmission is not negligible[48-50].

The application of HCV risk scores can be influenced by programmatic considerations and variations in context. This may necessitate prioritizing specific diagnostic metrics, such as sensitivity over specificity. The approach presented here demonstrates an inherent flexibility of the score, allowing adjustments to address specific programmatic needs, as illustrated by the analysis using different cut-off points (Table 3). However, it is critical to acknowledge the inherent trade-offs between diagnostic metrics. Optimizing one metric, such as sensitivity, will inevitably impact others, like specificity. Therefore, careful consideration is essential to align the score's cut-off with the specific programmatic context and its corresponding needs.

This study has limitations. For ease of use in primary healthcare and more broadly by the public, a risk score has to be simple. Accordingly, it cannot fully represent the complex epidemiology of HCV infection, such as interactions among risk factors. This risk score was derived for Egypt, which may not benefit from this risk score, given that this country has opted for mass testing of its entire population[17]. Derivation of a risk score typically requires at least one round of a population-based survey, ideally at the national level, but many countries may not have such survey data to be able to easily derive a risk score. The risk score was derived for a high-burden country, and utility of this approach still needs to be demonstrated for countries with low HCV prevalence. Nonetheless, this approach may prove to have higher utility in countries with low HCV prevalence than in countries with high HCV prevalence, as HCV epidemiology shows a

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

clearer hierarchy in infection exposure risk in countries with concentrated HCV epidemics compared to those with generalized HCV epidemics[46].

Conclusions

An HCV risk score can be derived using only one round of a population-based survey and offers an effective, simple, non-invasive strategy to identify carriers of HCV infection and to link them to testing and treatment, at low cost. This public health tool can be implemented and used for prioritizing populations for interventions with minimal logistical complexity and cost, especially in resource-limited countries.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgements The authors are grateful for the administrative support of Ms. Adona Canlas.

Ethics approval The study used anonymized publicly available data accessed through the Demographic and Health Survey program; therefore, ethical approval or collection of informed consent was not required.

Contributors REK conducted the data analyses with HC and NN. REK, NN, and LJA co-wrote the first draft of the article. LJA conceived and led the design of the study, analyses, and drafting the article. All authors contributed to drafting and revising the manuscript. All authors have read and approved the final manuscript.

Funding This work was supported by the National Priorities Research Program (NPRP) [grant number 12S-0216-190094] from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors. The authors are also grateful for infrastructure support provided by the Biostatistics, Epidemiology, and Biomathematics Research Core at Weill Cornell Medicine-Qatar.

Competing interests None declared.

Patient consent for publication Not required.

Data availability statement All data analyzed in this study can be accessed through application to the DHS Program at https://dhsprogram.com/ or by contacting archive@dhsprogram.com.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1 Results of the multivariable regression analyses to derive the Egypt Hepatitis C Risk Score using data from EDHS 2008 and EDHS 2014

	EDHS 2008				EDHS 2014			
	aOR*† (95% CI)	p-value	β‡	Risk score§	aOR*† (95% CI)	p-value	β‡	Risk score§
Sex								
Female	1.00		Ref	0	1.00		Ref	0
Male	1.52 (1.34-1.73)	<0.001	0.42	4	1.62 (1.40-1.87)	<0.001	0.48	5
Age group (years)								
15-19	1.00			0	1.00			
20-24	1.23 (0.89-1.69)	0.213	0.20	2	3.30 (1.88-5.81)	<0.001	1.19	12
25-29	1.60 (1.15-2.23)	0.005	0.47	5	4.50 (2.60-7.79)	<0.001	1.51	15
30-34	3.21 (2.35-4.39)	<0.001	1.17	12	7.41 (4.35-12.65)	<0.001	2.00	20
35-39	3.89 (2.84-5.34)	<0.001	1.36	14	8.74 (5.13-14.88)	<0.001	2.17	22
40-44	7.36 (5.47-9.90)	<0.001	1.99	20	13.03 (7.79-21.81)	<0.001	2.57	26
45-49	10.34 (7.71-13.85)	<0.001	2.34	23	19.23 (11.66-31.69)	<0.001	2.96	30
50-54	16.43 (12.29-21.96)	<0.001	2.80	28	41.11 (25.05-67.46)	<0.001	3.71	37
55-59	17.05 (12.50-23.26)	<0.001	2.84	28	55.31 (33.59-91.06)	<0.001	4.01	40
Type of place of residence								
Urban	1.00		Ref	0	1.00		Ref	0
Rural	2.34 (2.0-2.7)	<0.001	0.85	9	2.15 (1.78-2.59)	<0.001	0.76	8

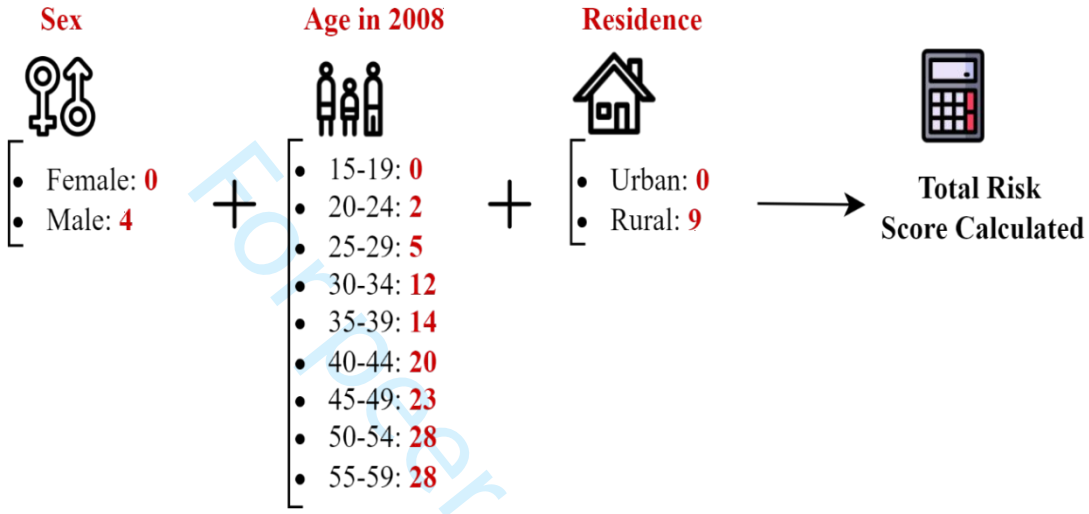
Abbreviations: aOR, Adjusted odds ratio; CI, Confidence interval; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; Ref, Reference category.
*The analysis applied the EDHS sampling weights.
†The odds ratio was adjusted for sex, age, and type of place of residence.
‡ β-coefficients were based on the multivariable regression analysis.
§The risk score was calculated by multiplying the β coefficient by 10 and then rounding the result to the nearest integer.

For peer review only

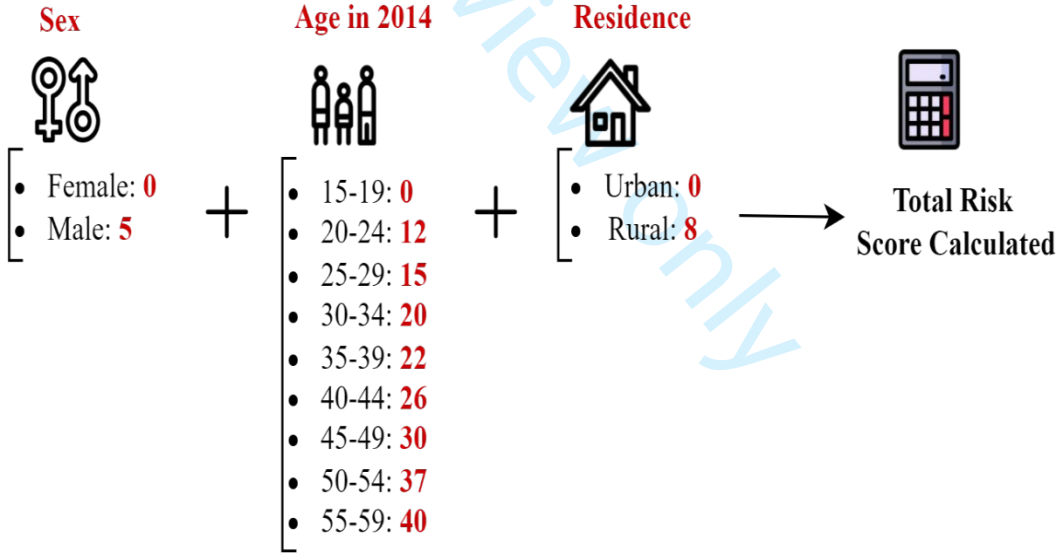
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Mathematical formula of the derived Egypt Hepatitis C Risk Score. A) Egypt Hepatitis C Risk Score using the EDHS 2008. B) Egypt Hepatitis C Risk Score using the EDHS 2014.

A.



B.

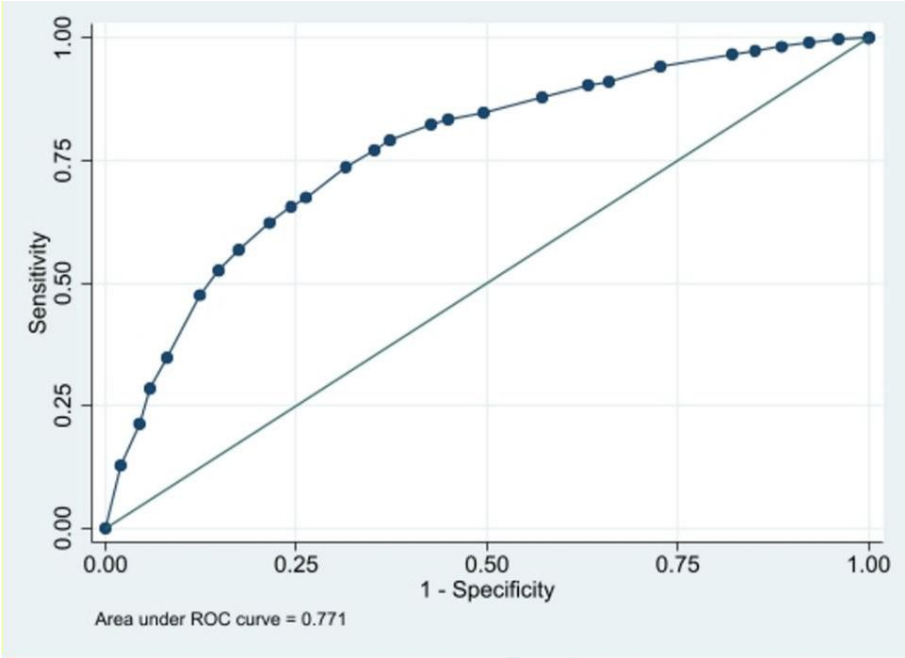


Abbreviations: EDHS, Egypt Demographic and Health Survey; ROC, receiver operating characteristics.
The two scores exhibited comparable structures and diagnostic performances, with minor differences attributed to sampling variation of the same population across two distinct rounds of the EDHS surveys. Details on the derivation of these scores are provided in the Methods and Results sections.

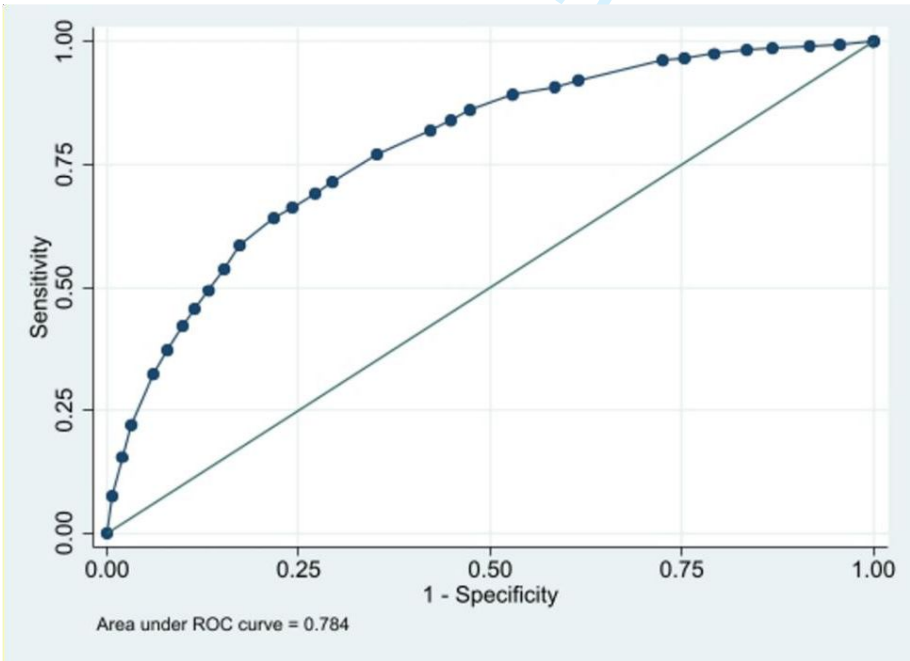
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 2. Diagnostic performance of the Egypt Hepatitis C Risk Score using the area under the (ROC) curve. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

A.



B.



Abbreviations: EDHS, Egypt Demographic and Health Survey; ROC, receiver operating characteristics.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2 Performance of the Egypt Hepatitis C Risk Score

	AUC (95% CI)	Risk score cut-off*	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportion needing testing (95% CI)
Derived risk scores							
Risk score derived using the EDHS 2008	0.77 (0.76-0.78)	22.0	73.7 (71.5-75.9)	68.5 (67.5-69.4)	27.8 (26.4-29.2)	94.1 (93.5-94.6)	37.2 (36.3-38.1)
Risk score derived using the EDHS 2014	0.78 (0.77- 0.80)	34.5	64.0 (61.5-66.6)	78.2 (77.5-78.9)	22.2 (20.9-23.5)	95.7 (95.4-96.1)	25.5 (24.9-26.2)
Validation of risk scores							
2008 risk score applied to the EDHS data 2014†	0.75 (0.74-0.77)	22.0	66.1 (63.5-68.6)	72.3 (71.5-73.1)	21.9 (20.6-23.2)	94.8 (94.3-95.2)	31.7 (30.9-32.5)
2014 risk score applied to the EDHS data 2008‡	0.76 (0.74-0.77)	33.5	70.0 (67.5-72.6)	70.0 (69.0-70.9)	24.7 (23.3-26.1)	94.3 (93.7-94.9)	34.6 (33.7-35.5)

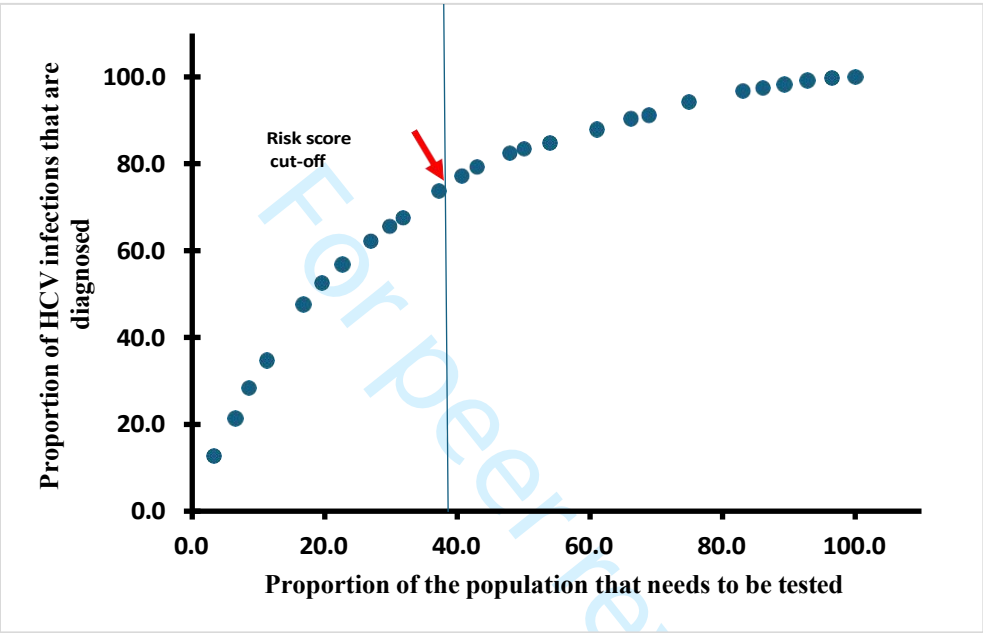
Abbreviations: AUC, Area under the curve; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; NPV, Negative predictive value; PPV, Positive predictive value.
*The optimal cut-off for the score was determined by maximizing the sum of the sensitivity and specificity.
†The risk score assumes the age of the individuals in 2008 in order to account for the age shift.
‡The risk score assumes the age of the individuals in 2014 in order to account for the age shift.

For peer review only

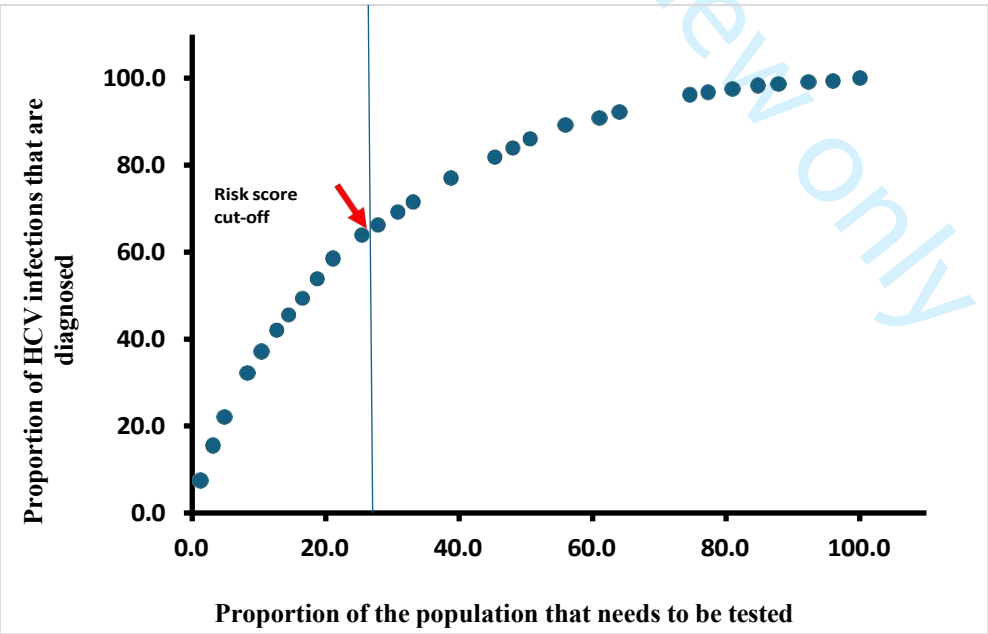
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 3. Proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections. The figure shows the effect of prioritization of testing for those with higher to lower risk score. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

A.



B.



For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3 Implications of selecting various score cutoffs on the different diagnostic metrics.

Risk score cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportion needing testing (95% CI)	Proportion that are diagnosed (95% CI)
EDHS 2008 Risk Score*						
≥2.0	99.8 (99.4-100.0)	3.9 (3.5-4.3)	14.6 (13.9-15.3)	99.2 (97.7-99.8)	96.5 (96.2-96.8)	99.8 (99.4-100.0)
≥4.0	99.1 (98.5-99.5)	7.8 (7.3-8.4)	15.0 (14.3-15.7)	98.2 (96.9-99.0)	92.8 (92.8-93.2)	99.1 (98.1-99.5)
≥5.0	98.2 (97.4-98.8)	11.4 (10.8-12.1)	15.4 (14.7-16.1)	97.5 (96.4-98.3)	89.3 (88.8-89.9)	98.2 (97.4-98.8)
≥6.0	97.5 (69.5-98.2)	14.8 (14.1-15.6)	15.8 (15.1-16.6)	97.3 (96.3-98.0)	86.1 (85.5-86.7)	97.5 (96.6-98.2)
≥9.0	96.8 (95.8-97.6)	17.9 (17.1-18.7)	16.2 (15.5-17.0)	97.1 (96.2-97.8)	83.2 (82.5-83.9)	96.8 (95.8-97.6)
≥11.0	94.2 (92.9-95.3)	27.3 (26.4-28.2)	17.6 (16.7-18.4)	96.6 (95.6-97.3)	74.9 (74.1-75.7)	94.2 (92.9-95.3)
≥12.0	91.2 (89.6-92.5)	34.0 (33.1-35.0)	18.5 (17.6-19.4)	95.9 (95.2-96.5)	68.9 (68.0-69.7)	91.2 (89.6-92.5)
≥13.0	90.5 (88.9-91.9)	36.8 (35.8-37.7)	19.0 (18.2-20.0)	95.9 (95.2-96.5)	66.3 (65.4-67.1)	90.5 (88.9-91.9)
≥14.0	87.9 (86.2-89.5)	42.8 (41.8-43.8)	20.2 (19.2-21.1)	95.6 (94.9-69.2)	61.0 (60.1-61.9)	87.9 (86.2-89.5)
≥15.0	84.8 (82.9-86.5)	50.5 (49.5-51.5)	22.1 (21.0-23.1)	95.3 (94.7-95.8)	54.0 (53.1-54.9)	84.8 (82.9-86.5)
≥16.0	83.5 (81.5-85.3)	55.1 (54.0-56.1)	23.74 (22.6-24.9)	95.3 (94.7-95.8)	50.1 (49.2-50.9)	83.5 (81.5-85.3)
≥18.0	82.4 (80.5-84.3)	57.3 (56.3-58.3)	24.1 (23.0-25.3)	95.2 (94.6-95.7)	47.9 (47.0-48.8)	82.4 (80.5-84.3)
≥20.0	79.1 (77.0-81.1)	62.7 (61.7-63.7)	25.8 (24.6-27.1)	94.8 (94.2-95.3)	43.0 (42.1-43.9)	79.1 (77.0-81.1)
≥21.0	77.1 (75.0-79.2)	64.7 (63.8-65.7)	26.5 (25.2-27.8)	94.5 (93.9-95.0)	40.8 (39.9-41.6)	77.2 (75.0-79.2)
≥23.0	73.7 (71.5-78.9)	68.5 (67.5-69.4)	27.8 (26.4-29.2)	94.1 (93.5-94.6)	37.2 (36.3-38.1)	73.7 (71.5-75.9)
≥24.0	67.5 (65.1-69.8)	73.7 (72.8-74.6)	29.7 (28.2-31.2)	93.3 (92.6-93.8)	31.8 (30.9-32.6)	67.5 (65.1-69.8)
≥25.0	65.6 (63.2-67.9)	75.7 (74.8-76.5)	30.7 (29.1-32.3)	93.0 (92.4-93.6)	29.8 (29.0-30.6)	65.6 (63.2-67.9)
≥27.0	62.3 (59.8-64.7)	78.5 (77.7-73.3)	32.3 (30.6-34.0)	92.7 (92.1-93.2)	27.0 (26.2-27.8)	62.3 (59.8-64.7)
≥28.0	56.7 (54.2-59.2)	82.6 (81.8-83.3)	34.9 (33.0-36.8)	92.1 (91.5-92.6)	22.7 (21.9-23.4)	56.7 (54.2-59.2)
≥29.0	52.5 (50.0-55.0)	85.1 (84.4-85.8)	36.7 (34.8-38.8)	91.6 (91.0-92.2)	19.6 (18.9-20.3)	52.5 (50.0-55.0)
≥32.0	47.6 (45.1-50.1)	87.6 (86.9-88.3)	38.7 (36.5-40.9)	91.0 (90.4-91.6)	16.8 (16.2-17.5)	47.6 (45.1-50.1)
≥33.0	34.7 (32.3-37.1)	92.0 (91.4-92.5)	41.5 (38.8-44.2)	89.5 (88.9-90.1)	11.3 (10.7-11.9)	34.7 (32.3-37.1)
≥36.0	28.4 (26.2-30.7)	94.2 (93.7-94.7)	44.6 (41.4-47.7)	88.9 (88.3-89.5)	8.6 (8.1-9.1)	28.4 (26.2-30.7)
≥37.0	21.3 (19.3-23.4)	95.6 (95.2-96.0)	44.3 (40.7-47.9)	88.1 (87.4-88.7)	6.5 (6.1-7.0)	21.3 (19.3-23.4)
≥41.0	12.7 (11.1-14.5)	98.1 (97.8-98.3)	51.8 (46.7-56.9)	87.2 (86.6-87.9)	3.3 (3.0-3.6)	12.7 (11.1-14.5)
EDHS 2014 Risk Score†						
≥5.0	99.4 (98.8-99.7)	4.3 (4.0-4.7)	9.1 (8.7-9.6)	98.6 (97.3-99.4)	96.0 (95.7-96.3)	99.4 (98.8-99.7)
≥8.0	99.2 (98.5-99.6)	8.3 (7.9-8.9)	9.5 (9.0-10.0)	99.0 (98.3-99.5)	92.3 (91.9-92.7)	99.2 (98.5-99.6)
≥12.0	98.7 (97.9-99.2)	13.2 (12.6-13.7)	9.9 (9.4-10.4)	99.0 (98.5-99.4)	87.9 (87.3-88.3)	98.7 (97.9-99.2)
≥13.0	98.3 (97.5-98.9)	16.5 (15.9-17.1)	10.2 (9.7-10.8)	99.0 (98.5-99.4)	84.7 (84.2-85.3)	98.3 (97.5-98.9)
≥15.0	97.5 (96.6-98.3)	20.7 (20.1-21.4)	10.7 (10.1-11.2)	98.9 (98.4-99.2)	80.9 (80.3-81.5)	97.5 (96.6-98.3)
≥17.0	96.7 (65.6-97.5)	24.7 (24.0-25.4)	11.1 (10.5-11.6)	98.7 (98.3-99.1)	77.2 (76.5-77.8)	96.7 (95.6-97.5)
≥20.0	96.2 (95.0-97.1)	27.5 (26.7-28.2)	11.4 (10.8-12.0)	98.7 (98.3-99.0)	74.5 (73.9-75.2)	96.2 (95.0-97.1)
≥22.0	92.2 (90.7-93.6)	38.5 (37.7-39.2)	12.7 (12.0-13.3)	98.1 (97.7-98.4)	64.0 (63.3-64.7)	92.2 (90.7-93.4)
≥23.0	90.8 (89.2-92.3)	41.5 (40.7-42.4)	13.1 (12.4-13.8)	97.9 (97.5-98.2)	61.0 (60.3-61.8)	90.8 (89.2-92.3)
≥25.0	89.2 (87.4-90.7)	47.1 (46.3-47.9)	14.0 (13.3-14.8)	97.8 (97.5-98.2)	55.9 (55.2-56.7)	89.2 (87.4-90.7)
≥26.0	86.1 (84.1-87.8)	52.6 (51.8-53.4)	15.0 (14.2-15.8)	97.5 (97.1-97.8)	50.6 (49.9-51.4)	86.1 (84.1-87.8)
≥27.0	83.9 (81.9-85.8)	55.1 (54.3-55.9)	15.3 (14.5-16.2)	97.3 (96.9-97.6)	48.1 (47.3-48.9)	83.9 (81.3-85.8)
≥28.0	81.9 (79.8-84.0)	57.8 (57.0-58.6)	15.8 (15.0-16.7)	97.1 (96.7-97.4)	45.5 (44.7-46.2)	81.9 (79.8-83.9)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

≥30.0	77.1 (74.8-79.2)	64.8 (64.0-65.6)	17.5 (16.6-18.5)	96.7 (96.3-97.0)	38.8 (38.0-39.5)	77.1 (74.8-79.2)
≥31.0	71.4 (69.0-73.8)	70.6 (69.8-71.3)	19.0 (18.0-20.1)	96.2 (95.8-96.6)	33.1 (32.3-33.8)	71.4 (69.0-73.8)
≥33.0	69.1 (66.7-71.5)	72.8 (72.1-73.5)	19.8 (18.7-20.9)	96.1 (95.7-96.4)	30.8 (30.1-31.5)	69.1 (66.7-71.5)
≥34.0	66.2 (63.7-68.7)	75.9 (75.2-76.6)	21.0 (19.8-22.2)	95.9 (95.5-96.2)	27.8 (27.1-28.5)	66.2 (63.7-68.7)
≥35.0	64.1 (61.5-66.6)	78.2 (77.5-78.9)	22.2 (20.9-23.5)	95.7 (95.4-96.1)	25.5 (24.9-26.2)	64.1 (61.5-66.6)
≥37.0	58.5 (55.9-61.1)	82.7 (82.0-83.3)	24.6 (23.2-26.1)	95.4 (95.0-95.7)	21.1 (20.5-21.7)	58.5 (55.9-61.1)
≥38.0	53.7 (51.1-56.3)	84.7 (84.1-85.2)	25.3 (23.8-26.9)	95.0 (94.6-95.3)	18.7 (18.1-19.3)	53.7 (51.1-56.3)
≥39.0	49.3 (46.7-52.0)	86.7 (86.1-87.2)	26.4 (24.7-28.1)	94.5 (94.2-95.0)	16.5 (15.9-17.1)	49.3 (46.7-52.0)
≥40.0	45.6 (43.0-48.2)	88.6 (88.1-89.1)	27.9 (26.1-29.8)	94.4 (94.0-94.8)	14.4 (13.1-15.0)	45.6 (43.0-48.2)
≥42.0	42.1 (39.5-44.7)	90.1 (89.6-90.6)	29.2 (27.3-31.3)	94.1 (93.7-94.5)	12.7 (12.2-13.2)	42.1 (39.5-44.7)
≥43.0	37.2 (34.6-39.7)	92.1 (91.7-92.6)	31.4 (29.1-33.6)	93.8 (93.4-94.2)	10.4 (9.9-10.8)	37.2 (34.6-39.7)
≥45.0	32.3 (29.8-34.8)	94.0 (93.6-94.4)	34.3 (31.7-36.9)	93.5 (93.1-93.9)	8.3 (7.9-8.7)	32.3 (29.8-34.8)
≥48.0	22.0 (19.9-24.3)	96.8 (96.5-97.1)	40.0 (36.6-43.6)	92.8 (92.2-93.2)	4.8 (4.5-5.1)	22.0 (19.9-24.3)
≥50.0	15.4 (13.6-17.4)	98.1 (97.8-98.3)	43.7 (39.3-48.2)	92.3 (91.9-92.7)	3.1 (2.8-3.4)	15.4 (13.6-17.4)
≥53.0	7.4 (6.1-8.9)	99.3 (99.2-99.4)	51.5 (44.4-58.5)	91.8 (91.4-92.3)	1.3 (1.1-1.4)	7.4 (6.1-8.9)

Abbreviations: AUC, Area under the curve; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; NPV, Negative predictive value; PPV, Positive predictive value.
*The AUC for the derived risk score using the EDHS 2008 was 0.77 (95% CI: 0.76-0.78).
†The AUC for the derived risk score using the EDHS 2014 was 0.78 (95% CI: 0.77-0.80).

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

References

1. Stanaway, J.D., et al., *The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013.* The Lancet, 2016. **388**(10049): p. 1081-1088.
2. World Health Organization, *Combating hepatitis B and C to reach elimination by 2030: advocacy brief*. 2016, World Health Organization.
3. Lauer, G.M. and B.D. Walker, *Hepatitis C virus infection.* New England journal of medicine, 2001. **345**(1): p. 41-52.
4. World Health Organization. *Hepatitis C.* 2021 27 July 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/hepatitis-c.
5. Calvaruso, V. and A. Craxì, *Hepatic benefits of HCV cure.* Journal of Hepatology, 2020. **73**(6): p. 1548-1556.
6. Ayoub, H.H. and L.J. Abu-Raddad, *Impact of treatment on hepatitis C virus transmission and incidence in Egypt: A case for treatment as prevention.* J Viral Hepat, 2017. **24**(6): p. 486-495.
7. Ayoub, H.H. and L.J. Abu-Raddad, *Treatment as prevention for hepatitis C virus in Pakistan: mathematical modelling projections.* BMJ Open, 2019. **9**(5): p. e026600.
8. World Health Organization, *Global health sector strategy on viral hepatitis 2016-2021. Towards ending viral hepatitis*. 2016, World Health Organization.
9. World Health Organization, *Global hepatitis report 2017*. 2017, World Health Organization.
10. World Health Organization, *Epidemiology of hepatitis C virus in the WHO Eastern Mediterranean Region: implications for strategic action*. 2020.
11. Chemaitelly, H., et al., *The epidemiology of hepatitis C virus in Afghanistan: systematic review and meta-analysis.* International Journal of Infectious Diseases, 2015. **40**: p. 54-63.
12. Mohamoud, Y.A., S. Riome, and L.J. Abu-Raddad, *Epidemiology of hepatitis C virus in the Arabian Gulf countries: Systematic review and meta-analysis of prevalence.* International Journal of Infectious Diseases, 2016. **46**: p. 116-125.
13. Chemaitelly, H., K. Chaabna, and L.J. Abu-Raddad, *The Epidemiology of Hepatitis C Virus in the Fertile Crescent: Systematic Review and Meta-Analysis.* PLOS ONE, 2015. **10**(8): p. e0135281.
14. Fadlalla, F.A., et al., *The Epidemiology of Hepatitis C Virus in the Maghreb Region: Systematic Review and Meta-Analyses.* PloS one, 2015. **10**(3): p. e0121873.
15. Chaabna, K., S.P. Kouyoumjian, and L.J. Abu-Raddad, *Hepatitis C virus epidemiology in Djibouti, Somalia, Sudan, and Yemen: systematic review and meta-analysis.* PloS one, 2016. **11**(2): p. e0149966.
16. Mahmud, S., V. Akbarzadeh, and L.J. Abu-Raddad, *The epidemiology of hepatitis C virus in Iran: Systematic review and meta-analyses.* Sci Rep, 2018. **8**(1): p. 150.
17. Waked, I., et al., *Screening and Treatment Program to Eliminate Hepatitis C in Egypt.* N Engl J Med, 2020. **382**(12): p. 1166-1174.
18. Al Kanaani, Z., et al., *The epidemiology of hepatitis C virus in Pakistan: systematic review and meta-analyses.* R Soc Open Sci, 2018. **5**(4): p. 180257.
19. Ayoub, H.H., Z. Al Kanaani, and L.J. Abu-Raddad, *Characterizing the temporal evolution of the hepatitis C virus epidemic in Pakistan.* J Viral Hepat, 2018. **25**(6): p. 670-679.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

20. Mahmud, S., Z. Al Kanaani, and L.J. Abu-Raddad, *Characterization of the hepatitis C virus epidemic in Pakistan.* BMC Infect Dis, 2019. **19**(1): p. 809.
21. Noble, D., et al., *Risk models and scores for type 2 diabetes: systematic review.* Bmj, 2011. **343**.
22. Collins, G.S., et al., *Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting.* BMC medicine, 2011. **9**(1): p. 1-14.
23. Brown, N., et al., *Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review.* Diabetes research and clinical practice, 2012. **98**(3): p. 369-385.
24. Awad, S.F., et al., *A diabetes risk score for Qatar utilizing a novel mathematical modeling approach to identify individuals at high risk for diabetes.* Scientific reports, 2021. **11**(1): p. 1-10.
25. El-Zanaty, F. and A. Way, *Egypt Demographic and Health Survey 2008. Cairo, Egypt Ministry of Health, El Zanaty and Associates, and Macro International.* 2009.
26. El-Zanaty, F., *Egypt Health Issue Survey 2015. Cairo, Egypt and Rockville, Maryland, USA: Ministry of Health and Population and ICF International.* 2015.
27. Rutstein, S.O. and G. Rojas, *Guide to DHS statistics.* Calverton, MD: ORC Macro, 2006. **38**.
28. MEASURE DHS. *The DHS Program Demographic and Health Surveys.* . Available from: https://dhsprogram.com/.
29. Ayoub, H.H., et al., *Characterizing the historical role of parenteral antischistosomal therapy in hepatitis C virus transmission in Egypt.* International journal of epidemiology, 2020. **49**(3): p. 798-809.
30. Mohamoud, Y.A., et al., *The epidemiology of hepatitis C virus in Egypt: a systematic review and data synthesis.* BMC infectious diseases, 2013. **13**(1): p. 1-21.
31. Kouyoumjian, S.P., H. Chemaitelly, and L.J. Abu-Raddad, *Characterizing hepatitis C virus epidemiology in Egypt: systematic reviews, meta-analyses, and meta-regressions.* Scientific reports, 2018. **8**(1): p. 1-17.
32. Jovanovic, M.R., et al., *Does the Strategy of Risk Group Testing for Hepatitis C Hit the Target?* Frontiers in Pharmacology, 2017. **8**.
33. Smart, A., et al., *Identification of Risk Factors for Testing of Hepatitis C in Non-Birth Cohort Patients: Is Universal Screening Necessary?* J Addict Med, 2021. **15**(2): p. 109-112.
34. Jordan, A.E. and D.C. Perlman, *The Shift in Emphasis From Risk-Based to Age-Based Hepatitis C Virus (HCV) Testing in the US Tends to Remove Injection Drug Use From Discourse on HCV.* Subst Use Misuse, 2017. **52**(3): p. 340-350.
35. Qureshi, H., et al., *Prevalence of hepatitis B and C viral infections in pakistan: Findings of a national survey appealing for effective prevention and control measures. [French].* Eastern Mediterranean Health Journal, 2010. **16**(SUPPL.): p. S15-23.
36. Daw, M.A. and A. El-Bouzedi, *Prevalence of hepatitis B and hepatitis C infection in Libya: results from a national population based survey.* BMC Infect Dis, 2014. **14**: p. 17.
37. Mahmud, S., et al., *Individual-level key associations and modes of exposure for hepatitis C virus infection in the Middle East and North Africa: a systematic synthesis.* Ann Epidemiol, 2018. **28**(7): p. 452-461.
38. Harfouche, M., et al., *Hepatitis C virus viremic rate in the Middle East and North Africa: Systematic synthesis, meta-analyses, and meta-regressions.* PLoS One, 2017. **12**(10): p.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

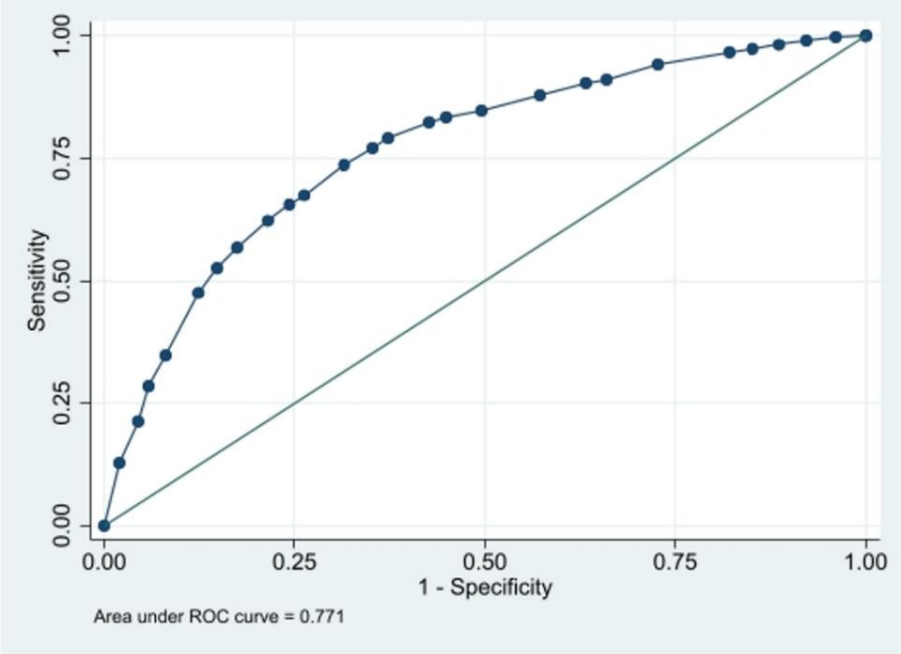
e0187177.
39. Harfouche, M., et al., *Epidemiology of hepatitis C virus among hemodialysis patients in the Middle East and North Africa: systematic syntheses, meta-analyses, and meta-regressions.* Epidemiol Infect, 2017. **145**(15): p. 3243-3263.
40. Heijnen, M., G.R. Mumtaz, and L.J. Abu-Raddad, *Status of HIV and hepatitis C virus infections among prisoners in the Middle East and North Africa: review and synthesis.* J Int AIDS Soc, 2016. **19**(1): p. 20873.
41. Mahmud, S., et al., *Hepatitis C Virus Infection in Populations With Liver-Related Diseases in the Middle East and North Africa.* Hepatol Commun, 2020. **4**(4): p. 577-587.
42. Mahmud, S., et al., *Characterizing trends and associations for hepatitis C virus antibody prevalence in the Middle East and North Africa: meta-regression analyses.* Sci Rep, 2022. **12**(1): p. 20637.
43. Mahmud, S., et al., *The status of hepatitis C virus infection among people who inject drugs in the Middle East and North Africa.* Addiction, 2020. **115**(7): p. 1244-1262.
44. Mohamoud, Y.A., F.D. Miller, and L.J. Abu-Raddad, *Potential for human immunodeficiency virus parenteral transmission in the Middle East and North Africa: an analysis using hepatitis C virus as a proxy biomarker.* World J Gastroenterol, 2014. **20**(36): p. 12734-52.
45. Mahmud, S., et al., *Key associations for hepatitis C virus genotypes in the Middle East and North Africa.* J Med Virol, 2020. **92**(3): p. 386-393.
46. Chemaitelly, H., et al., *Who to Test for Hepatitis C Virus in the Middle East and North Africa?: Pooled Analyses of 2,500 Prevalence Measures, Including 49 Million Tests.* Hepatol Commun, 2019. **3**(3): p. 325-339.
47. Mumtaz, G.R., et al., *Status of the HIV epidemic in key populations in the Middle East and north Africa: knowns and unknowns.* Lancet HIV, 2022. **9**(7): p. e506-e516.
48. Benova, L., S.F. Awad, and L.J. Abu-Raddad, *Estimate of vertical transmission of Hepatitis C virus in Pakistan in 2007 and 2012 birth cohorts.* J Viral Hepat, 2017. **24**(12): p. 1177-1183.
49. Benova, L., et al., *Estimation of hepatitis C virus infections resulting from vertical transmission in Egypt.* Hepatology, 2015. **61**(3): p. 834-42.
50. Benova, L., et al., *Vertical transmission of hepatitis C virus: systematic review and meta-analysis.* Clin Infect Dis, 2014. **59**(6): p. 765-73.

For peer review only

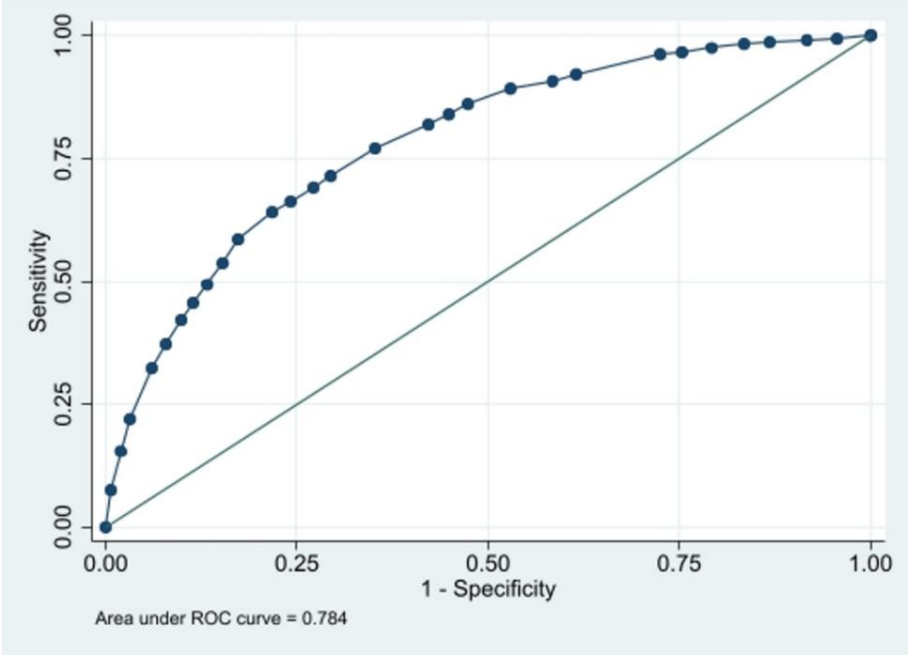
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Diagnostic performance of the Egypt Hepatitis C Risk Score using the area under the (ROC) curve. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

A.



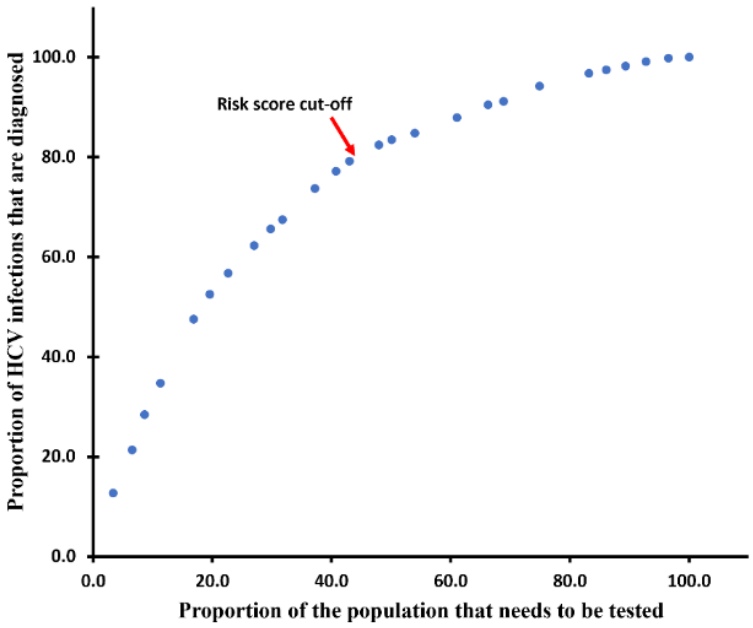
B.



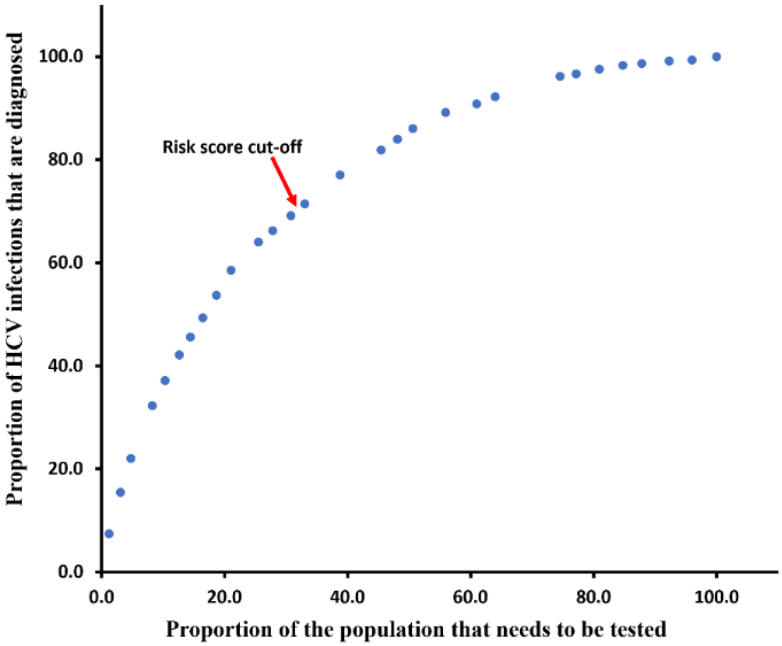
Abbreviations: EDHS, Egypt Demographic and Health Survey; ROC, receiver operating characteristics.

Figure 2. Proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections. The figure shows the effect of prioritization of testing for those with higher to lower risk score. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

A.



B.



Abbreviations: EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus.

Table S1 Characteristics of individuals tested for HCV antibodies in the EDHS 2008 and 2014

Characteristics	EDHS 2008			EDHS 2014		
	Total tested (%)	HCV antibody positive (proportion %)	p-value	Total tested (%)	HCV antibody positive (proportion %)	p-value
No	11,126	1,571		26,047	1,456	
Sex						
Female	6,052 (54.4)	711 (11.8)	<0.001	13,707 (52.6)	660 (4.8)	<0.001
Male	5,074 (45.6)	860 (17.0)		12,340 (47.4)	796 (6.4)	
Age group (years)						
1-4	-	-	-	3,282 (12.6)	10 (0.3)	<0.001
5-9	-	-	-	3,601 (13.8)	10 (0.3)	
10-14	-	-	-	3,161 (12.1)	23 (0.7)	
15-19	2,000 (18.0)	82 (4.1)	<0.001	2,568 (9.9)	30 (1.2)	
20-24	1,837 (16.5)	91 (5.0)		1,976 (7.6)	54 (2.7)	
25-29	1,520 (13.7)	92 (6.1)		2,358 (9.0)	88 (2.7)	
30-34	1,244 (11.2)	133 (10.7)		2,076 (8.0)	114 (5.5)	
35-39	1,141 (10.3)	147 (12.9)		1,853 (7.1)	130 (7.0)	
40-44	1,069 (9.6)	238 (22.3)		1,468 (5.6)	146 (10.0)	
45-49	939 (8.4)	275 (29.3)		1,380 (5.3)	208 (15.1)	
50-54	728 (6.5)	272 (37.4)		1,334 (5.1)	337 (25.3)	
55-59	648 (5.8)	241 (37.2)		990 (3.8)	306 (30.9)	
Type of place of residence						
Urban	4,448 (40.0)	442 (9.9)	<0.001	11,955 (45.9)	546 (4.6)	<0.001
Rural	6,678 (60.0)	1,129 (16.9)		14,092 (54.1)	910 (6.5)	

Abbreviations: EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus

For peer review only

Table S2 STROBE checklist for cross-sectional studies

	Item No	Recommendation	Main Text Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction
Methods			
Study design	4	Present key elements of study design early in the paper	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score')
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods ('Egypt Demographic and Health Surveys')
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	Methods ('Egypt Demographic and Health Surveys')
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') & Table S1
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') & Table 1& Table 3 & Tables S2-S3 & Figures 1-2 & Figure S2
Bias	9	Describe any efforts to address potential sources of bias	Methods ('Risk score derivation'& 'Performance and validation of the risk score') & Table 2 & Figures 1-2
Study size	10	Explain how the study size was arrived at	Not applicable, see Methods ('Egypt Demographic and Health Surveys')
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') & Tables 1-3 & Table S3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Methods ('Risk score derivation'& 'Performance and validation of the risk score')
		(*b*) Describe any methods used to examine subgroups and interactions	Not applicable, see Methods ('Risk score derivation'& 'Performance and validation of the risk score')
		(*c*) Explain how missing data were addressed	Not applicable, see Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation')
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	Not applicable
		(*e*) Describe any sensitivity analyses	Not applicable
Results			
Participants	13	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	Table 1 & Tables S1 & Table S3
Descriptive data	14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 1 & Table S1 & Table S3 & Figure S1
		(b) Indicate number of participants with missing data for each variable of interest	Not applicable, see Methods ('Egypt Demographic and Health Surveys')
Outcome data	15	Report numbers of outcome events or summary measures	Results, Figures 1-2, Figure S1, Tables 1-3 & Table S1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 1 & Table S3

For peer review only

		(*b*) Report category boundaries when continuous variables were categorized	Tables 1-3 & Table S1 & Table S3
		© If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Tables 2-3 & Table S3 & Figure S1
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Results & Table 2 & Figure 2
Discussion			
Key results	18	Summarise key results with reference to study objectives	Discussion, paragraphs 1-3
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion, paragraph 9
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion, paragraphs 4-6
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussion, paragraphs 7-8
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Acknowledgements

For peer review only

Table S3 Results of univariable and multivariable regression analyses to derive the Egypt Hepatitis C Risk Score using data from EDHS 2008 and EDHS 2014

Characteristics	EDHS 2008						EDHS 2014					
	OR* (95% CI)	p-value	aOR*† (95% CI)	p-value	β‡	Risk score§	OR* (95% CI)	p-value	aOR*† (95% CI)	p-value	β‡	Risk score§
Sex												
Female	1.00		1.00		Ref	0	1.00		1.00		Ref	0
Male	1.52 (1.35-1.70)	<0.001	1.52 (1.34-1.73)	<0.001	0.42	4	1.60 (1.41-1.83)	<0.001	1.62 (1.40-1.87)	<0.001	0.48	5
Age group (years)												
15-19	1.00		1.00			0	1.00		1.00			
20-24	1.19 (0.86-1.64)	0.282	1.23 (0.89-1.69)	0.213	0.20	2	3.18 (1.82-5.58)	<0.001	3.30 (1.88-5.81)	<0.001	1.19	12
25-29	1.52 (1.09-2.11)	0.014	1.60 (1.15-2.23)	0.005	0.47	5	4.38 (2.54-7.55)	<0.001	4.50 (2.60-7.79)	<0.001	1.51	15
30-34	3.09 (2.27-4.21)	<0.001	3.21 (2.35-4.39)	<0.001	1.17	12	7.33 (4.29-12.49)	<0.001	7.41 (4.35-12.65)	<0.001	2.00	20
35-39	3.69 (2.70-5.06)	<0.001	3.89 (2.84-5.34)	<0.001	1.36	14	8.56 (5.04-14.52)	<0.001	8.74 (5.13-14.88)	<0.001	2.17	22
40-44	6.91 (5.15-9.26)	<0.001	7.36 (5.47-9.90)	<0.001	1.99	20	12.54 (7.51-20.95)	<0.001	13.03 (7.79-21.81)	<0.001	2.57	26
45-49	9.30 (6.95-12.27)	<0.001	10.34 (7.71-13.85)	<0.001	2.34	23	18.64 (11.33-30.66)	<0.001	19.23 (11.66-31.69)	<0.001	2.96	30
50-54	14.36 (10.78-19.13)	<0.001	16.43 (12.29-21.96)	<0.001	2.80	28	37.25 (22.75-60.98)	<0.001	41.11 (25.05-67.46)	<0.001	3.71	37
55-59	15.09 (11.13-20.45)	<0.001	17.05 (12.50-23.26)	<0.001	2.84	28	49.26 (30.06-80.72)	<0.001	55.31 (33.59-91.06)	<0.001	4.01	40
Type of place of residence												
Urban	1.00		1.00		Ref	0	1.00		1.00		Ref	0
Rural	1.91 (1.66-2.19)	<0.001	2.34 (2.0-2.7)	<0.001	0.85	9	1.73 (1.46-2.1)	<0.001	2.15 (1.78-2.59)	<0.001	0.76	8

Abbreviations: aOR, Adjusted odds ratio; CI, Confidence interval; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; OR, Odds ratio; Ref, Reference category.
*The analysis applied the EDHS sampling weights.
†The odds ratio was adjusted for sex, age, and type of place of residence.
‡ β-coefficients were based on the multivariable regression analysis.
§The risk score was calculated by multiplying the β coefficient by 10 and then rounding the result to the nearest integer.

For peer review only

BMJ Open

Hepatitis C Risk Score as a Tool to Identify Individuals with HCV Infection: A Demonstration and Cross-Sectional Epidemiological Study in Egypt

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-085506.R1
Article Type:	Original research
Date Submitted by the Author:	02-Jun-2024
Complete List of Authors:	El-Khoury, Rayane; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group Chemaitelly, Hiam; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group; Weill Cornell Medicine - Qatar, World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis Alaama, Ahmed S.; World Health Organisation Regional Office for the Eastern Mediterranean, Department of Communicable Diseases Hermez, Joumana G.; World Health Organisation Regional Office for the Eastern Mediterranean, Department of Communicable Diseases Nagelkerke, Nico; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group Abu-Raddad, Laith; Weill Cornell Medicine - Qatar, Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar; Weill Cornell Medicine - Qatar, World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Infectious diseases
Keywords:	EPIDEMIOLOGY, Risk Factors, Public health < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

ORIGINAL RESEARCH

Hepatitis C Risk Score as a Tool to Identify Individuals with HCV Infection: A Demonstration and Cross-Sectional Epidemiological Study in Egypt

Rayane El-Khoury[1,2], Hiam Chemaitelly[1,2,3], Ahmed S. Alaama[4], Joumana G. Hermez[4], and Nico Nagelkerke[1], Laith J. Abu-Raddad[1,2,3,5,6*]

[1]Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar, Cornell University, Qatar Foundation - Education City, Doha, Qatar
[2]World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis, Weill Cornell Medicine–Qatar, Cornell University, Qatar Foundation – Education City, Doha, Qatar
[3]Department of Population Health Sciences, Weill Cornell Medicine, Cornell University, New York, New York, USA
[4]Department of Communicable Diseases, HIV/Hepatitis/STIs Unit, World Health Organization Regional Office for the Eastern Mediterranean, Cairo, Egypt
[5]Department of Public Health, College of Health Sciences, Member of QU Health, Qatar University, Doha, Qatar
[6]College of Health and Life Sciences, Hamad bin Khalifa University, Doha, Qatar

Word count: Abstract: 279 words, Main text: 3,388 words.
Number of tables: 3.
Number of figures: 3.
Running head: Hepatitis C virus risk score in Egypt.
Funding: The Qatar National Research Fund (NPRP 12S-0216-190094).

***Correspondence to:** Professor Laith J. Abu-Raddad, Infectious Disease Epidemiology Group, World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis, Weill Cornell Medicine - Qatar, Qatar Foundation - Education City, P.O. Box 24144, Doha, Qatar. Telephone: +(974) 4492-8321. Fax: +(974) 4492-8333. E-mail: lja2002@qatar-med.cornell.edu.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ABSTRACT

Objectives: Hepatitis C virus (HCV) infection poses a global health challenge. By the end of 2021, the World Health Organization estimated that less than a quarter of global HCV infections had been diagnosed. There is a need for a public health tool that can facilitate the identification of people with HCV infection and link them to testing and treatment, and that can be customized for each country.

Methods: We derived and validated a risk score to identify people with HCV in Egypt and demonstrated its utility. Utilizing data from the 2008 and 2014 Egypt Demographic and Health Surveys, two risk scores were constructed through multivariable logistic regression analysis. A range of diagnostic metrics was then calculated to evaluate the performance of these scores.

Results: The 2008 and 2014 risk scores exhibited similar dependencies on sex, age, and type of place of residence. Both risk scores demonstrated high and similar areas under the curve of 0.77 (95% CI: 0.76-0.78) and 0.78 (95% CI: 0.77-0.80), respectively. For the 2008 Risk Score, sensitivity was 73.7% (95% CI: 71.5-75.9%), specificity was 68.5% (95% CI: 67.5-69.4%), positive predictive value (PPV) was 27.8% (95% CI: 26.4-29.2%), and negative predictive value (NPV) was 94.1% (95% CI: 93.5-94.6%). For the 2014 Risk Score, sensitivity was 64.0% (95% CI: 61.5-66.6%), specificity was 78.2% (95% CI: 77.5-78.9%), PPV was 22.2% (95% CI: 20.9-23.5%), and NPV was 95.7% (95% CI: 95.4-96.1%). Each score was validated by applying it to a different survey database than the one used to derive it.

Conclusions: Implementation of HCV risk scores is an effective strategy to identify carriers of HCV infection and to link them to testing and treatment at low cost to national programs.

Keywords: Hepatitis C virus, viral hepatitis, risk score, Egypt, Demographic Health Survey.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study:

- This study derived a risk score that provides a non-invasive and easily administered method to identify hepatitis C virus carriers and link them to treatment and care.
- The risk score was based on and validated using two rounds of population-based, high-quality national surveys in Egypt.
- The derivation of the risk score used only a few variables and thus may not adequately capture the complex epidemiology of hepatitis C virus infection.
- The derived risk score is specific to Egypt and may not be applicable to populations in other countries.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

INTRODUCTION

Hepatitis C virus (HCV) infection is a global public health challenge[1, 2] and a major cause of morbidity and mortality, resulting in liver cancer, fibrosis, and cirrhosis[3]. By end of 2021, the World Health Organization (WHO) estimated that 58 million people were infected with HCV, but only 15 million of them were diagnosed and only 9 million received treatment[4]. Direct-acting antivirals (DAA) offer highly effective treatment to cure this infection and to prevent progression toward severe forms of liver disease[5], as well as an opportunity to reduce HCV transmission through treatment as prevention[6, 7]. Accordingly, the WHO has set a global target to eliminate HCV infection as a public health problem by 2030[2, 8].

While DAAs are becoming accessible globally, it has been challenging to identify carriers of this infection so as to treat them, especially in the Middle East and North Africa (MENA), the region most affected by HCV infection and where most people with HCV infection remain undiagnosed[9, 10]. Limited resources have made it challenging for viral hepatitis programs to find low-cost and cost-effective approaches to identify people with HCV. While mass testing and treatment programs may be relevant in high prevalence countries, other countries have relatively low HCV prevalence making such programs less cost-effective[10-16]. While low-cost point-of-care tests (POCs) have been beneficial in some countries, such as Egypt[17], they remain relatively expensive for countries like Pakistan, which bear a substantial share of the global burden[18-20]. There is a need for a public health tool that can assist in identifying persons potentially living with HCV, to link them to testing and treatment.

One such tool is the use of risk scores to identify individuals potentially living with HCV. A risk score comprises a small set of simple questions that can be used to assess the likelihood that an individual has a specific health condition[21-24], in this case, HCV infection. Such risk scores

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

have proven influential as public health tools for a range of health conditions, such as diabetes[21-24].

In this study, we demonstrate the application of this public health tool for HCV infection in Egypt, aiming to illustrate the public health value and practical utility of developing HCV risk scores in various countries. The risk score derived here is not intended for universal application across diverse settings; it is specifically designed for Egypt. However, the concept and analytical approach can be adapted to other countries by considering the local HCV epidemiology to determine the relevant factors and their respective weights for inclusion in a score tailored to each specific context.

METHODS

Egypt Demographic and Health Surveys

The Egypt Demographic and Health Survey (EDHS) is a national survey that collected data pertaining to the health and demographics of a nationally representative sample of the resident population of Egypt, including HCV infection[25, 26]. The EDHS that included HCV biomarkers was conducted in 2008 and 2014 and used rigorous sampling methods[27]. Details on study design, data collection, and laboratory methods can be found in El-Zanaty et al.[25, 26].

HCV antibody testing was done using a third generation enzyme-linked immunosorbent assay (ELISA), the Enzyme Immunoassay Adlatis EIAgen HCV Ab test (Adaltis Inc., Montreal, Canada)[25, 26]. All samples that were positive in the ELISA assay and 5% of the negative samples were then retested using a more specific assay, the chemiluminescent microplate immunoassay (CMIA ARCHITECT plus i1000SR, Abbott Diagnostic, USA)[25, 26]. If a sample was positive in both the ELISA and the CMIA testing, it was also tested for current

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

active infection, using real-time, reverse-transcription polymerase chain reaction (RT-qPCR) testing to detect HCV ribonucleic acid (RNA)[25, 26]. Samples were further retested for internal and external quality assurance[25, 26]. Here we restrict our analyses to the HCV antibody results.

Data from the EDHS 2008 and EDHS 2014 were downloaded with permission from Measure DHS[28]. The data can be accessed through application to the DHS Program at https://dhsprogram.com. For purposes of this study, the EDHS individual database was merged with the HCV biomarker database, based on established guidelines for managing DHS data[27]. All individuals with results for HCV antibody testing were included in the analysis.

Patient and Public Involvement

None.

Risk score derivation

Associations of HCV antibody positivity (seropositivity) with a priori variables that are easy to evaluate in a primary healthcare setting, and that can be included in a risk score, were investigated. These variables included sex (male versus female), age (5-year age strata), and type of place of residence (urban versus rural). Frequency distributions were generated to describe demographic and clinical profiles of tested individuals.

Chi-square tests and univariable logistic regression were implemented to investigate associations. Participants younger than 15 years of age were excluded as this age group was not included in the EDHS 2008 and has low HCV prevalence (Table S1)[6, 29-31]. Odds ratios (ORs), 95% confidence intervals (CIs), and p-values were reported. Covariates with p-values ≤0.1 in univariable regression analysis were considered possibly associated with HCV

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

seropositivity. These were included in the multivariable analysis for estimation of adjusted ORs (AORs) and associated 95% CIs and p-values. No other forward or backward elimination for variable selection was used. Covariates with p-values ≤0.05 in the multivariable model were considered predictors of HCV seropositivity. Univariable and multivariable analyses were adjusted for sampling weights.

A risk score was constructed based on the β-coefficients obtained from the multivariable regression model. β-coefficients were multiplied by a factor of 10 and then rounded to the nearest integer. The total risk score was calculated by adding the individual scores. To keep the score simple enough for use in primary healthcare and other general population settings, we did not consider any interaction terms.

Performance and validation of the risk score

A receiver operating characteristics (ROC) curve was plotted to investigate the performance of the risk score in predicting HCV seropositivity at different score cut-offs. A larger area under the curve (AUC), also called the c-index, indicates better performance of the risk score. The cut-off for the score was determined by maximizing the sum of the sensitivity and specificity. Sensitivity is the probability that the risk score will yield a positive diagnosis in a subject who is truly HCV antibody-positive. Specificity is the probability that the risk score will yield a negative diagnosis in a subject who is truly HCV antibody-negative.

Performance of the risk score was also investigated by estimating the positive predictive value (PPV) and the negative predictive value (NPV) of the risk score. PPV is the probability that a subject with a positive diagnosis per the risk score is truly HCV antibody-positive. NPV is the probability that a subject with a negative diagnosis per the risk score is truly HCV antibody-negative. The proportion of subjects who have scores greater than or equal to the cut-off of the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

risk score was estimated to determine the proportion of individuals that need to be biochemically tested for HCV antibodies.

To validate the performance of the EDHS 2008 risk score, it was applied to the EDHS 2014 data, providing an independent validation with a dataset different from the one used for its derivation. Performance diagnostics were subsequently assessed. Given the pronounced cohort effect in the epidemiology of HCV infection in Egypt[6, 29-31], the age variable was adjusted to reflect the 6-year interval between the surveys. For example, individuals who were 11 years old in 2008 would have been 17 years old at the time of the second survey in 2014. The same approach was also used to validate the EDHS 2014 risk score—it was applied to the EDHS 2008 database and performance diagnostics were assessed.

While the cut-off for the score was determined by maximizing the sum of sensitivity and specificity, this cut-off can be adjusted as needed from a programmatic standpoint to optimize a specific diagnostic metric, such as sensitivity instead of specificity. To illustrate this flexibility, an additional analysis was incorporated featuring a variety of score cut-offs, resulting in diverse values of sensitivity, specificity, PPV, and NPV. Such additional analysis enables program managers and readers to discern the trade-offs among these diagnostic metrics and observe the implications of selecting an alternative programmatic approach, such as prioritizing the optimization of sensitivity over specificity.

Analyses were conducted in Stata version 16.1 (Stata Corporation, College Station, TX, USA). The study was reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Table S2).

RESULTS

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

In the 2008 EDHS, 11,126 individuals 15-59 years of age were tested, of whom 1,571 were antibody-positive[25]. The 2014 EDHS included children 1-14 years of age in addition to adults 15-59 years of age[26]. In this latter survey, 26,047 individuals were tested of whom 1,456 were antibody-positive[26].

Characteristics of individuals who were tested for HCV antibodies and the proportion of each population stratum that was HCV antibody-positive are shown in Table S1 for both of the EDHS surveys. Results of both surveys were consistent, taking into account the age shift in the national cohort with the passage of 6 years between the EDHS 2008 and EDHS 2014.

HCV seropositivity was strongly associated with sex, age, and place of residence in both national surveys (Table 1 and Table S3). Male sex and rural residence were associated with higher seropositivity. Seropositivity increased rapidly with age.

The 2008 and 2014 Egypt Hepatitis C Risk Scores derived using the EDHS 2008 and EDHS 2014 data, respectively, are shown in Figure 1. The 2008 Risk Score had a range of 0-41. The 2014 Risk Score had a range of 0-53. Both showed similar dependence on sex, age, and type of place of residence. Both demonstrated high and similar AUCs (Figure 2). The AUC was 0.77 (95% CI: 0.76-0.78) for the 2008 Risk Score and 0.78 (95% CI: 0.77-0.80) for the 2014 Risk Score. The highest sum of sensitivity and specificity was obtained at a score cut-off value of 22 for the 2008 Risk Score and at a cut-off of 34.5 for the 2014 Risk Score.

For the 2008 Risk Score, sensitivity was 73.7% (95% CI: 71.5-75.9%), specificity was 68.5% (95% CI: 67.5-69.4%), PPV was 27.8% (95% CI: 26.4-29.2%), and NPV was 94.1% (95% CI: 93.5-94.6%) (Table 2). For the 2014 Risk Score, sensitivity was 64.0% (95% CI: 61.5-66.6%), specificity was 78.2% (95% CI: 77.5-78.9%), PPV was 22.2% (95% CI: 20.9-23.5%), and NPV was 95.7% (95% CI: 95.4-96.1%). The proportion of the population 15-59 years of age that

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

needed to be biochemically tested for HCV antibodies was 37.2% (95% CI: 36.3-38.1%) using the 2008 Risk Score and 25.5% (95% CI: 24.9-26.2%) using the 2014 Risk Score. Of all people with HCV in the EDHS samples, application of this score would have diagnosed (that is identified; sensitivity) 73.7% (95% CI: 71.5-75.9%) and 64.0% (95% CI: 61.5-66.6%) of all these persons in samples of the EDHS 2008 and 2014, respectively.

When the 2008 Risk Score was applied to the EDHS 2014 data, the AUC was 0.75 (95% CI: 0.74-0.77), the sensitivity was 66.1% (95% CI: 63.5-68.6%), and the specificity was 72.3% (95% CI: 71.5-73.1%) (Table 2). These performance indicators were similar to the original performance indicators generated using the EDHS 2008 data, as well as to the performance indicators of the 2014 Risk Score on the EDHS 2014 data. Therefore, this application validates this risk score. A similar outcome was found when the 2014 Risk Score was applied to the EDHS 2008 data, also providing a validation of the 2014 risk score.

Figure 3 displays the proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections, using each of EDHS 2008 and EDHS 2014 data. The figure shows the effect of prioritization of testing for those with higher to lower risk score. This provides a demonstration of the utility of using the risk score: a large proportion of HCV infections can be diagnosed by testing only a small proportion of the population. It is most efficient programmatically to start testing individuals with the highest risk score and progressively moving on to those with lower and lower risk scores. As testing is expanded to those with low risk scores, the yield in identifying more HCV infections is very limited.

Table 3 illustrates the implications of selecting various score cut-offs, providing insight into the trade-offs among different diagnostic metrics, as well as the proportion of the population

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

requiring biochemical testing and the proportion of all individuals with HCV identified through the application of this score. For instance, by enhancing the specificity of the risk score, the PPV increases, and the proportion of the population necessitating testing decreases. This reduction in testing requirements helps alleviate costs and streamline the logistics of the test-and-treat program. However, this enhanced program efficiency comes at the expense of lower NPV and sensitivity, implying a smaller proportion of individuals with HCV in the population being identified through the risk score.

DISCUSSION

We demonstrated that a risk score that consists of few simple questions that are easy to evaluate in a primary healthcare setting or implemented through a website or an app that helps persons identify their risk of being HCV infected, provides an effective and non-invasive public health tool to identify carriers of HCV infection and to link them to testing and treatment. Biochemical testing methods to identify people with HCV are invasive and time-consuming and require human and financial resources, as well as complex logistics, making them less scalable, particularly in resource-limited settings. In contrast, initial screening using a risk score can be easily administered or self-administered, is non-invasive, and requires minimal resources and logistics. Therefore, HCV risk scores can be an indispensable strategy for the global response to attain the target of HCV elimination as a public health problem by 2030.

While the concept of a risk score shares similarities with risk-based testing, which has been implemented in some countries, predominantly in higher-income nations[32-34], the risk score approach transcends mere risk-based testing. It enables a broader application across various settings and situations and can significantly contribute to raising awareness of the infection among the general population. The risk score approach represents a tool that addresses several

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

public health needs simultaneously, extending the application of risk-based testing beyond conventional healthcare settings. Moreover, it entails minimal costs and logistics, making it feasible even in resource-limited settings.

Remarkably, the risk score, comprising three simple questions, demonstrated considerable diagnostic accuracy, as evidenced by the values of various diagnostic metrics, including AUC, sensitivity, specificity, PPV, and NPV. Of particular note is the high NPV, ensuring that a negative result is highly unlikely to be a false negative, thereby obviating the need for individuals with a negative outcome using the score to undergo testing for HCV antibodies. The score also identified 73.7% and 64.0% of all HCV infections in the EDHS 2008 and EDHS 2014 samples, respectively. Thus, the score fulfills its objective of facilitating the efficient identification of individuals with HCV infection while minimizing the necessity for conducting biochemical testing. This underscores the value of this approach in identifying as many people with HCV as early as possible and initiating treatment before progression to serious clinical disease.

This approach was demonstrated for Egypt, considering the availability of two EDHS surveys to derive and validate the score. The two scores exhibited comparable structures and diagnostic performances, with minor differences attributed to sampling variation of the same population across two distinct rounds of the EDHS surveys. Each score was validated by applying it to a database other than the one used to derive it. The latter application yielded a diagnostic performance that was comparable to the original diagnostic performance against the database used to originate it. This highlights how a single national survey for HCV infection may be sufficient to develop an effective risk score for this infection, and that can become an integral component of the national response to eliminate HCV infection.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The approach demonstrated in this study can be applied in other countries, including those in the MENA region. In countries where nationally representative population-based surveys have been conducted, these surveys can serve as the basis for deriving the risk score, as was done in this study. However, only three MENA countries—Egypt, Libya, and Pakistan—have conducted such surveys[25, 26, 35, 36]. For countries where such surveys are not available[10-16], the risk score can still be derived using data from available regional surveys. Alternatively, if regional surveys are not available, the effects of risk factors for infection can be pooled, either in terms of odds ratios or relative risks, using data from analytical studies[37]. These effects can also be derived from meta-regression analyses applied to all available HCV prevalence studies for each country[38-45].

While this study focused on demonstrating the utility of this concept as a public health tool, actual application of this approach to different countries can be enhanced for even higher diagnostic accuracy. One extension could be adding more variables to the score in a manner tailored to the local epidemiology of each country. For instance, province or city of birth and/or current residence, prior exposure to an HCV mode of transmission[37], or history of HCV infection in the family, could be added, among others. Given that the risk of exposure to HCV infection varies immensely by at-risk population type and shows a distinctive hierarchy[46], an additional component to the score could be to integrate the at-risk population type as a variable[41, 46], thereby further enhancing the diagnostic accuracy of the score. Testing strategies, therefore, could be highly efficient in identifying people with HCV at a modest cost.

However, caution must be exercised to prevent the creation of stigma associated with HCV infection or the use of an HCV risk score. For instance, it may not be feasible to include questions about stigmatized behaviors in the MENA context, such as injecting drug use or

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

specific sexual practices, when the score is applied in general population settings like primary healthcare. However, such questions may be appropriate in other settings, such as voluntary counseling and testing (VCT) centers or outreach efforts by community organizations working with the most at-risk populations[47]. It is important also for the risk score to factor community acceptance in its design and implementation, ensuring it addresses the specific needs of certain groups, such as women of childbearing age in contexts where the risk of HCV vertical transmission is not negligible[48-50].

The application of HCV risk scores can be influenced by programmatic considerations and variations in context. This may necessitate prioritizing specific diagnostic metrics, such as sensitivity over specificity. The approach presented here demonstrates an inherent flexibility of the score, allowing adjustments to address specific programmatic needs, as illustrated by the analysis using different cut-off points (Table 3). However, it is critical to acknowledge the inherent trade-offs between diagnostic metrics. Optimizing one metric, such as sensitivity, will inevitably impact others, like specificity. Therefore, careful consideration is essential to align the score's cut-off with the specific programmatic context and its corresponding needs.

This study has limitations. For ease of use in primary healthcare and more broadly by the public, a risk score has to be simple. Accordingly, it cannot fully represent the complex epidemiology of HCV infection, such as interactions among risk factors. This risk score was derived for Egypt, which may not benefit from this risk score, given that this country has opted for mass testing of its entire population[17]. Derivation of a risk score typically requires at least one round of a population-based survey, ideally at the national level, but many countries may not have such survey data to be able to easily derive a risk score. The risk score was derived for a high-burden country, and utility of this approach still needs to be demonstrated for countries with low HCV

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

prevalence. Nonetheless, this approach may prove to have higher utility in countries with low HCV prevalence than in countries with high HCV prevalence, as HCV epidemiology shows a clearer hierarchy in infection exposure risk in countries with concentrated HCV epidemics compared to those with generalized HCV epidemics[46].

Conclusions

An HCV risk score can be derived using only one round of a population-based survey and offers an effective, simple, non-invasive strategy to identify carriers of HCV infection and to link them to testing and treatment, at low cost. This public health tool can be implemented and used for prioritizing populations for interventions with minimal logistical complexity and cost, especially in resource-limited countries.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgements The authors are grateful for the administrative support of Ms. Adona Canlas.

Ethics approval The study used anonymized publicly available data accessed through the Demographic and Health Survey program; therefore, ethical approval or collection of informed consent was not required.

Contributors REK conducted the data analyses with HC and NN. REK, NN, and LJA co-wrote the first draft of the article. LJA conceived and led the design of the study, analyses, and drafting the article. All authors contributed to drafting and revising the manuscript. All authors have read and approved the final manuscript.

Funding This work was supported by the National Priorities Research Program (NPRP) [grant number 12S-0216-190094] from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors. The authors are also grateful for infrastructure support provided by the Biostatistics, Epidemiology, and Biomathematics Research Core at Weill Cornell Medicine-Qatar.

Competing interests None declared.

Patient consent for publication Not required.

Data availability statement All data analyzed in this study can be accessed through application to the DHS Program at https://dhsprogram.com/ or by contacting archive@dhsprogram.com.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Mathematical formula of the derived Egypt Hepatitis C Risk Score. A) Egypt Hepatitis C Risk Score using the EDHS 2008. B) Egypt Hepatitis C Risk Score using the EDHS 2014.

Figure 2. Diagnostic performance of the Egypt Hepatitis C Risk Score using the area under the (ROC) curve. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

Figure 3. Proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections. The figure shows the effect of prioritization of testing for those with higher to lower risk score. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1 Results of the multivariable regression analyses to derive the Egypt Hepatitis C Risk Score using data from EDHS 2008 and EDHS 2014

	EDHS 2008				EDHS 2014			
	aOR*† (95% CI)	p-value	β‡	Risk score§	aOR*† (95% CI)	p-value	β‡	Risk score§
Sex								
Female	1.00		Ref	0	1.00		Ref	0
Male	1.52 (1.34-1.73)	<0.001	0.42	4	1.62 (1.40-1.87)	<0.001	0.48	5
Age group (years)								
15-19	1.00			0	1.00			
20-24	1.23 (0.89-1.69)	0.213	0.20	2	3.30 (1.88-5.81)	<0.001	1.19	12
25-29	1.60 (1.15-2.23)	0.005	0.47	5	4.50 (2.60-7.79)	<0.001	1.51	15
30-34	3.21 (2.35-4.39)	<0.001	1.17	12	7.41 (4.35-12.65)	<0.001	2.00	20
35-39	3.89 (2.84-5.34)	<0.001	1.36	14	8.74 (5.13-14.88)	<0.001	2.17	22
40-44	7.36 (5.47-9.90)	<0.001	1.99	20	13.03 (7.79-21.81)	<0.001	2.57	26
45-49	10.34 (7.71-13.85)	<0.001	2.34	23	19.23 (11.66-31.69)	<0.001	2.96	30
50-54	16.43 (12.29-21.96)	<0.001	2.80	28	41.11 (25.05-67.46)	<0.001	3.71	37
55-59	17.05 (12.50-23.26)	<0.001	2.84	28	55.31 (33.59-91.06)	<0.001	4.01	40
Type of place of residence								
Urban	1.00		Ref	0	1.00		Ref	0
Rural	2.34 (2.0-2.7)	<0.001	0.85	9	2.15 (1.78-2.59)	<0.001	0.76	8

Abbreviations: aOR, Adjusted odds ratio; CI, Confidence interval; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; Ref, Reference category.
*The analysis applied the EDHS sampling weights.
†The odds ratio was adjusted for sex, age, and type of place of residence.
‡ β-coefficients were based on the multivariable regression analysis.
§The risk score was calculated by multiplying the β coefficient by 10 and then rounding the result to the nearest integer.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2 Performance of the Egypt Hepatitis C Risk Score

	AUC (95% CI)	Risk score cut-off*	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportion needing testing (95% CI)
Derived risk scores							
Risk score derived using the EDHS 2008	0.77 (0.76-0.78)	22.0	73.7 (71.5-75.9)	68.5 (67.5-69.4)	27.8 (26.4-29.2)	94.1 (93.5-94.6)	37.2 (36.3-38.1)
Risk score derived using the EDHS 2014	0.78 (0.77- 0.80)	34.5	64.0 (61.5-66.6)	78.2 (77.5-78.9)	22.2 (20.9-23.5)	95.7 (95.4-96.1)	25.5 (24.9-26.2)
Validation of risk scores							
2008 risk score applied to the EDHS data 2014†	0.75 (0.74-0.77)	22.0	66.1 (63.5-68.6)	72.3 (71.5-73.1)	21.9 (20.6-23.2)	94.8 (94.3-95.2)	31.7 (30.9-32.5)
2014 risk score applied to the EDHS data 2008‡	0.76 (0.74-0.77)	33.5	70.0 (67.5-72.6)	70.0 (69.0-70.9)	24.7 (23.3-26.1)	94.3 (93.7-94.9)	34.6 (33.7-35.5)

Abbreviations: AUC, Area under the curve; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; NPV, Negative predictive value; PPV, Positive predictive value.
*The optimal cut-off for the score was determined by maximizing the sum of the sensitivity and specificity.
†The risk score assumes the age of the individuals in 2008 in order to account for the age shift.
‡The risk score assumes the age of the individuals in 2014 in order to account for the age shift.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3 Implications of selecting various score cutoffs on the different diagnostic metrics.

Risk score cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Proportion needing testing (95% CI)	Proportion that are diagnosed (95% CI)
EDHS 2008 Risk Score*						
≥2.0	99.8 (99.4-100.0)	3.9 (3.5-4.3)	14.6 (13.9-15.3)	99.2 (97.7-99.8)	96.5 (96.2-96.8)	99.8 (99.4-100.0)
≥4.0	99.1 (98.5-99.5)	7.8 (7.3-8.4)	15.0 (14.3-15.7)	98.2 (96.9-99.0)	92.8 (92.8-93.2)	99.1 (98.1-99.5)
≥5.0	98.2 (97.4-98.8)	11.4 (10.8-12.1)	15.4 (14.7-16.1)	97.5 (96.4-98.3)	89.3 (88.8-89.9)	98.2 (97.4-98.8)
≥6.0	97.5 (69.5-98.2)	14.8 (14.1-15.6)	15.8 (15.1-16.6)	97.3 (96.3-98.0)	86.1 (85.5-86.7)	97.5 (96.6-98.2)
≥9.0	96.8 (95.8-97.6)	17.9 (17.1-18.7)	16.2 (15.5-17.0)	97.1 (96.2-97.8)	83.2 (82.5-83.9)	96.8 (95.8-97.6)
≥11.0	94.2 (92.9-95.3)	27.3 (26.4-28.2)	17.6 (16.7-18.4)	96.6 (95.6-97.3)	74.9 (74.1-75.7)	94.2 (92.9-95.3)
≥12.0	91.2 (89.6-92.5)	34.0 (33.1-35.0)	18.5 (17.6-19.4)	95.9 (95.2-96.5)	68.9 (68.0-69.7)	91.2 (89.6-92.5)
≥13.0	90.5 (88.9-91.9)	36.8 (35.8-37.7)	19.0 (18.2-20.0)	95.9 (95.2-96.5)	66.3 (65.4-67.1)	90.5 (88.9-91.9)
≥14.0	87.9 (86.2-89.5)	42.8 (41.8-43.8)	20.2 (19.2-21.1)	95.6 (94.9-69.2)	61.0 (60.1-61.9)	87.9 (86.2-89.5)
≥15.0	84.8 (82.9-86.5)	50.5 (49.5-51.5)	22.1 (21.0-23.1)	95.3 (94.7-95.8)	54.0 (53.1-54.9)	84.8 (82.9-86.5)
≥16.0	83.5 (81.5-85.3)	55.1 (54.0-56.1)	23.74 (22.6-24.9)	95.3 (94.7-95.8)	50.1 (49.2-50.9)	83.5 (81.5-85.3)
≥18.0	82.4 (80.5-84.3)	57.3 (56.3-58.3)	24.1 (23.0-25.3)	95.2 (94.6-95.7)	47.9 (47.0-48.8)	82.4 (80.5-84.3)
≥20.0	79.1 (77.0-81.1)	62.7 (61.7-63.7)	25.8 (24.6-27.1)	94.8 (94.2-95.3)	43.0 (42.1-43.9)	79.1 (77.0-81.1)
≥21.0	77.1 (75.0-79.2)	64.7 (63.8-65.7)	26.5 (25.2-27.8)	94.5 (93.9-95.0)	40.8 (39.9-41.6)	77.2 (75.0-79.2)
≥23.0	73.7 (71.5-78.9)	68.5 (67.5-69.4)	27.8 (26.4-29.2)	94.1 (93.5-94.6)	37.2 (36.3-38.1)	73.7 (71.5-75.9)
≥24.0	67.5 (65.1-69.8)	73.7 (72.8-74.6)	29.7 (28.2-31.2)	93.3 (92.6-93.8)	31.8 (30.9-32.6)	67.5 (65.1-69.8)
≥25.0	65.6 (63.2-67.9)	75.7 (74.8-76.5)	30.7 (29.1-32.3)	93.0 (92.4-93.6)	29.8 (29.0-30.6)	65.6 (63.2-67.9)
≥27.0	62.3 (59.8-64.7)	78.5 (77.7-73.3)	32.3 (30.6-34.0)	92.7 (92.1-93.2)	27.0 (26.2-27.8)	62.3 (59.8-64.7)
≥28.0	56.7 (54.2-59.2)	82.6 (81.8-83.3)	34.9 (33.0-36.8)	92.1 (91.5-92.6)	22.7 (21.9-23.4)	56.7 (54.2-59.2)
≥29.0	52.5 (50.0-55.0)	85.1 (84.4-85.8)	36.7 (34.8-38.8)	91.6 (91.0-92.2)	19.6 (18.9-20.3)	52.5 (50.0-55.0)
≥32.0	47.6 (45.1-50.1)	87.6 (86.9-88.3)	38.7 (36.5-40.9)	91.0 (90.4-91.6)	16.8 (16.2-17.5)	47.6 (45.1-50.1)
≥33.0	34.7 (32.3-37.1)	92.0 (91.4-92.5)	41.5 (38.8-44.2)	89.5 (88.9-90.1)	11.3 (10.7-11.9)	34.7 (32.3-37.1)
≥36.0	28.4 (26.2-30.7)	94.2 (93.7-94.7)	44.6 (41.4-47.7)	88.9 (88.3-89.5)	8.6 (8.1-9.1)	28.4 (26.2-30.7)
≥37.0	21.3 (19.3-23.4)	95.6 (95.2-96.0)	44.3 (40.7-47.9)	88.1 (87.4-88.7)	6.5 (6.1-7.0)	21.3 (19.3-23.4)
≥41.0	12.7 (11.1-14.5)	98.1 (97.8-98.3)	51.8 (46.7-56.9)	87.2 (86.6-87.9)	3.3 (3.0-3.6)	12.7 (11.1-14.5)
EDHS 2014 Risk Score†						
≥5.0	99.4 (98.8-99.7)	4.3 (4.0-4.7)	9.1 (8.7-9.6)	98.6 (97.3-99.4)	96.0 (95.7-96.3)	99.4 (98.8-99.7)
≥8.0	99.2 (98.5-99.6)	8.3 (7.9-8.9)	9.5 (9.0-10.0)	99.0 (98.3-99.5)	92.3 (91.9-92.7)	99.2 (98.5-99.6)
≥12.0	98.7 (97.9-99.2)	13.2 (12.6-13.7)	9.9 (9.4-10.4)	99.0 (98.5-99.4)	87.9 (87.3-88.3)	98.7 (97.9-99.2)
≥13.0	98.3 (97.5-98.9)	16.5 (15.9-17.1)	10.2 (9.7-10.8)	99.0 (98.5-99.4)	84.7 (84.2-85.3)	98.3 (97.5-98.9)
≥15.0	97.5 (96.6-98.3)	20.7 (20.1-21.4)	10.7 (10.1-11.2)	98.9 (98.4-99.2)	80.9 (80.3-81.5)	97.5 (96.6-98.3)
≥17.0	96.7 (65.6-97.5)	24.7 (24.0-25.4)	11.1 (10.5-11.6)	98.7 (98.3-99.1)	77.2 (76.5-77.8)	96.7 (95.6-97.5)
≥20.0	96.2 (95.0-97.1)	27.5 (26.7-28.2)	11.4 (10.8-12.0)	98.7 (98.3-99.0)	74.5 (73.9-75.2)	96.2 (95.0-97.1)
≥22.0	92.2 (90.7-93.6)	38.5 (37.7-39.2)	12.7 (12.0-13.3)	98.1 (97.7-98.4)	64.0 (63.3-64.7)	92.2 (90.7-93.4)
≥23.0	90.8 (89.2-92.3)	41.5 (40.7-42.4)	13.1 (12.4-13.8)	97.9 (97.5-98.2)	61.0 (60.3-61.8)	90.8 (89.2-92.3)
≥25.0	89.2 (87.4-90.7)	47.1 (46.3-47.9)	14.0 (13.3-14.8)	97.8 (97.5-98.2)	55.9 (55.2-56.7)	89.2 (87.4-90.7)
≥26.0	86.1 (84.1-87.8)	52.6 (51.8-53.4)	15.0 (14.2-15.8)	97.5 (97.1-97.8)	50.6 (49.9-51.4)	86.1 (84.1-87.8)
≥27.0	83.9 (81.9-85.8)	55.1 (54.3-55.9)	15.3 (14.5-16.2)	97.3 (96.9-97.6)	48.1 (47.3-48.9)	83.9 (81.3-85.8)
≥28.0	81.9 (79.8-84.0)	57.8 (57.0-58.6)	15.8 (15.0-16.7)	97.1 (96.7-97.4)	45.5 (44.7-46.2)	81.9 (79.8-83.9)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

≥30.0	77.1 (74.8-79.2)	64.8 (64.0-65.6)	17.5 (16.6-18.5)	96.7 (96.3-97.0)	38.8 (38.0-39.5)	77.1 (74.8-79.2)
≥31.0	71.4 (69.0-73.8)	70.6 (69.8-71.3)	19.0 (18.0-20.1)	96.2 (95.8-96.6)	33.1 (32.3-33.8)	71.4 (69.0-73.8)
≥33.0	69.1 (66.7-71.5)	72.8 (72.1-73.5)	19.8 (18.7-20.9)	96.1 (95.7-96.4)	30.8 (30.1-31.5)	69.1 (66.7-71.5)
≥34.0	66.2 (63.7-68.7)	75.9 (75.2-76.6)	21.0 (19.8-22.2)	95.9 (95.5-96.2)	27.8 (27.1-28.5)	66.2 (63.7-68.7)
≥35.0	64.1 (61.5-66.6)	78.2 (77.5-78.9)	22.2 (20.9-23.5)	95.7 (95.4-96.1)	25.5 (24.9-26.2)	64.1 (61.5-66.6)
≥37.0	58.5 (55.9-61.1)	82.7 (82.0-83.3)	24.6 (23.2-26.1)	95.4 (95.0-95.7)	21.1 (20.5-21.7)	58.5 (55.9-61.1)
≥38.0	53.7 (51.1-56.3)	84.7 (84.1-85.2)	25.3 (23.8-26.9)	95.0 (94.6-95.3)	18.7 (18.1-19.3)	53.7 (51.1-56.3)
≥39.0	49.3 (46.7-52.0)	86.7 (86.1-87.2)	26.4 (24.7-28.1)	94.5 (94.2-95.0)	16.5 (15.9-17.1)	49.3 (46.7-52.0)
≥40.0	45.6 (43.0-48.2)	88.6 (88.1-89.1)	27.9 (26.1-29.8)	94.4 (94.0-94.8)	14.4 (13.1-15.0)	45.6 (43.0-48.2)
≥42.0	42.1 (39.5-44.7)	90.1 (89.6-90.6)	29.2 (27.3-31.3)	94.1 (93.7-94.5)	12.7 (12.2-13.2)	42.1 (39.5-44.7)
≥43.0	37.2 (34.6-39.7)	92.1 (91.7-92.6)	31.4 (29.1-33.6)	93.8 (93.4-94.2)	10.4 (9.9-10.8)	37.2 (34.6-39.7)
≥45.0	32.3 (29.8-34.8)	94.0 (93.6-94.4)	34.3 (31.7-36.9)	93.5 (93.1-93.9)	8.3 (7.9-8.7)	32.3 (29.8-34.8)
≥48.0	22.0 (19.9-24.3)	96.8 (96.5-97.1)	40.0 (36.6-43.6)	92.8 (92.2-93.2)	4.8 (4.5-5.1)	22.0 (19.9-24.3)
≥50.0	15.4 (13.6-17.4)	98.1 (97.8-98.3)	43.7 (39.3-48.2)	92.3 (91.9-92.7)	3.1 (2.8-3.4)	15.4 (13.6-17.4)
≥53.0	7.4 (6.1-8.9)	99.3 (99.2-99.4)	51.5 (44.4-58.5)	91.8 (91.4-92.3)	1.3 (1.1-1.4)	7.4 (6.1-8.9)

Abbreviations: AUC, Area under the curve; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; NPV, Negative predictive value; PPV, Positive predictive value.
*The AUC for the derived risk score using the EDHS 2008 was 0.77 (95% CI: 0.76-0.78).
†The AUC for the derived risk score using the EDHS 2014 was 0.78 (95% CI: 0.77-0.80).

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

References

1. Stanaway, J.D., et al., *The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013.* The Lancet, 2016. **388**(10049): p. 1081-1088.
2. World Health Organization, *Combating hepatitis B and C to reach elimination by 2030: advocacy brief*. 2016, World Health Organization.
3. Lauer, G.M. and B.D. Walker, *Hepatitis C virus infection.* New England journal of medicine, 2001. **345**(1): p. 41-52.
4. World Health Organization. *Hepatitis C.* 2021, 27 July 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/hepatitis-c.
5. Calvaruso, V. and A. Craxì, *Hepatic benefits of HCV cure.* Journal of Hepatology, 2020. **73**(6): p. 1548-1556.
6. Ayoub, H.H. and L.J. Abu-Raddad, *Impact of treatment on hepatitis C virus transmission and incidence in Egypt: A case for treatment as prevention.* J Viral Hepat, 2017. **24**(6): p. 486-495.
7. Ayoub, H.H. and L.J. Abu-Raddad, *Treatment as prevention for hepatitis C virus in Pakistan: mathematical modelling projections.* BMJ Open, 2019. **9**(5): p. e026600.
8. World Health Organization, *Global health sector strategy on viral hepatitis 2016-2021. Towards ending viral hepatitis*. 2016, World Health Organization.
9. World Health Organization, *Global hepatitis report 2017*. 2017, World Health Organization.
10. World Health Organization, *Epidemiology of hepatitis C virus in the WHO Eastern Mediterranean Region: implications for strategic action*. 2020.
11. Chemaitelly, H., et al., *The epidemiology of hepatitis C virus in Afghanistan: systematic review and meta-analysis.* International Journal of Infectious Diseases, 2015. **40**: p. 54-63.
12. Mohamoud, Y.A., S. Riome, and L.J. Abu-Raddad, *Epidemiology of hepatitis C virus in the Arabian Gulf countries: Systematic review and meta-analysis of prevalence.* International Journal of Infectious Diseases, 2016. **46**: p. 116-125.
13. Chemaitelly, H., K. Chaabna, and L.J. Abu-Raddad, *The Epidemiology of Hepatitis C Virus in the Fertile Crescent: Systematic Review and Meta-Analysis.* PLOS ONE, 2015. **10**(8): p. e0135281.
14. Fadlalla, F.A., et al., *The Epidemiology of Hepatitis C Virus in the Maghreb Region: Systematic Review and Meta-Analyses.* PloS one, 2015. **10**(3): p. e0121873.
15. Chaabna, K., S.P. Kouyoumjian, and L.J. Abu-Raddad, *Hepatitis C virus epidemiology in Djibouti, Somalia, Sudan, and Yemen: systematic review and meta-analysis.* PloS one, 2016. **11**(2): p. e0149966.
16. Mahmud, S., V. Akbarzadeh, and L.J. Abu-Raddad, *The epidemiology of hepatitis C virus in Iran: Systematic review and meta-analyses.* Sci Rep, 2018. **8**(1): p. 150.
17. Waked, I., et al., *Screening and Treatment Program to Eliminate Hepatitis C in Egypt.* N Engl J Med, 2020. **382**(12): p. 1166-1174.
18. Al Kanaani, Z., et al., *The epidemiology of hepatitis C virus in Pakistan: systematic review and meta-analyses.* R Soc Open Sci, 2018. **5**(4): p. 180257.
19. Ayoub, H.H., Z. Al Kanaani, and L.J. Abu-Raddad, *Characterizing the temporal evolution of the hepatitis C virus epidemic in Pakistan.* J Viral Hepat, 2018. **25**(6): p. 670-679.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

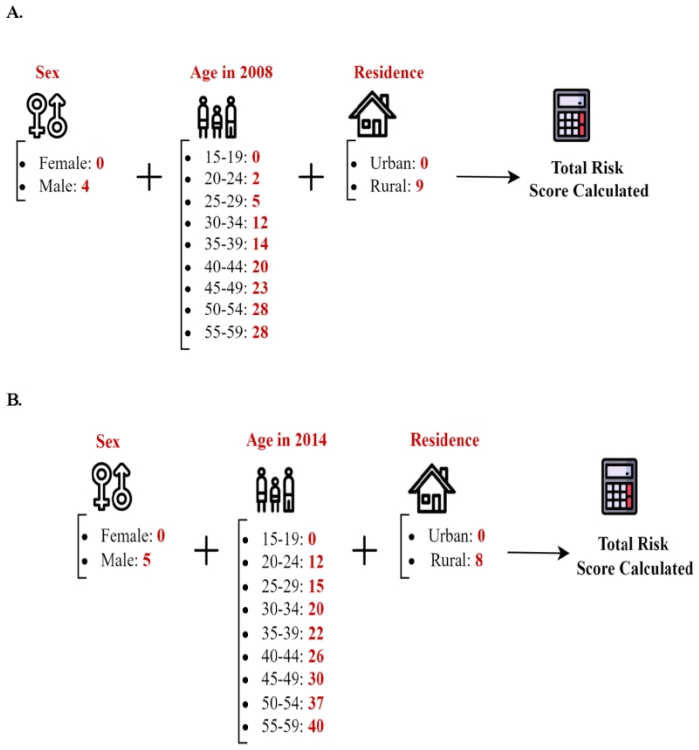
20. Mahmud, S., Z. Al Kanaani, and L.J. Abu-Raddad, *Characterization of the hepatitis C virus epidemic in Pakistan.* BMC Infect Dis, 2019. **19**(1): p. 809.
21. Noble, D., et al., *Risk models and scores for type 2 diabetes: systematic review.* Bmj, 2011. **343**.
22. Collins, G.S., et al., *Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting.* BMC medicine, 2011. **9**(1): p. 1-14.
23. Brown, N., et al., *Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review.* Diabetes research and clinical practice, 2012. **98**(3): p. 369-385.
24. Awad, S.F., et al., *A diabetes risk score for Qatar utilizing a novel mathematical modeling approach to identify individuals at high risk for diabetes.* Scientific reports, 2021. **11**(1): p. 1-10.
25. El-Zanaty, F. and A. Way, *Egypt Demographic and Health Survey 2008. Cairo, Egypt Ministry of Health, El Zanaty and Associates, and Macro International.* 2009.
26. El-Zanaty, F., *Egypt Health Issue Survey 2015. Cairo, Egypt and Rockville, Maryland, USA: Ministry of Health and Population and ICF International.* 2015.
27. Rutstein, S.O. and G. Rojas, *Guide to DHS statistics.* Calverton, MD: ORC Macro, 2006. **38**.
28. *[dataset]* MEASURE DHS. *The DHS Program Demographic and Health Surveys*. , July 2021; Available from: https://dhsprogram.com/.
29. Ayoub, H.H., et al., *Characterizing the historical role of parenteral antischistosomal therapy in hepatitis C virus transmission in Egypt.* International journal of epidemiology, 2020. **49**(3): p. 798-809.
30. Mohamoud, Y.A., et al., *The epidemiology of hepatitis C virus in Egypt: a systematic review and data synthesis.* BMC infectious diseases, 2013. **13**(1): p. 1-21.
31. Kouyoumjian, S.P., H. Chemaitelly, and L.J. Abu-Raddad, *Characterizing hepatitis C virus epidemiology in Egypt: systematic reviews, meta-analyses, and meta-regressions.* Scientific reports, 2018. **8**(1): p. 1-17.
32. Jovanovic, M.R., et al., *Does the Strategy of Risk Group Testing for Hepatitis C Hit the Target?* Frontiers in Pharmacology, 2017. **8**.
33. Smart, A., et al., *Identification of Risk Factors for Testing of Hepatitis C in Non-Birth Cohort Patients: Is Universal Screening Necessary?* J Addict Med, 2021. **15**(2): p. 109-112.
34. Jordan, A.E. and D.C. Perlman, *The Shift in Emphasis From Risk-Based to Age-Based Hepatitis C Virus (HCV) Testing in the US Tends to Remove Injection Drug Use From Discourse on HCV.* Subst Use Misuse, 2017. **52**(3): p. 340-350.
35. Qureshi, H., et al., *Prevalence of hepatitis B and C viral infections in pakistan: Findings of a national survey appealing for effective prevention and control measures. [French].* Eastern Mediterranean Health Journal, 2010. **16**(SUPPL.): p. S15-23.
36. Daw, M.A. and A. El-Bouzedi, *Prevalence of hepatitis B and hepatitis C infection in Libya: results from a national population based survey.* BMC Infect Dis, 2014. **14**: p. 17.
37. Mahmud, S., et al., *Individual-level key associations and modes of exposure for hepatitis C virus infection in the Middle East and North Africa: a systematic synthesis.* Ann Epidemiol, 2018. **28**(7): p. 452-461.
38. Harfouche, M., et al., *Hepatitis C virus viremic rate in the Middle East and North Africa: Systematic synthesis, meta-analyses, and meta-regressions.* PLoS One, 2017. **12**(10): p.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

e0187177.
39. Harfouche, M., et al., *Epidemiology of hepatitis C virus among hemodialysis patients in the Middle East and North Africa: systematic syntheses, meta-analyses, and meta-regressions.* Epidemiol Infect, 2017. **145**(15): p. 3243-3263.
40. Heijnen, M., G.R. Mumtaz, and L.J. Abu-Raddad, *Status of HIV and hepatitis C virus infections among prisoners in the Middle East and North Africa: review and synthesis.* J Int AIDS Soc, 2016. **19**(1): p. 20873.
41. Mahmud, S., et al., *Hepatitis C Virus Infection in Populations With Liver-Related Diseases in the Middle East and North Africa.* Hepatol Commun, 2020. **4**(4): p. 577-587.
42. Mahmud, S., et al., *Characterizing trends and associations for hepatitis C virus antibody prevalence in the Middle East and North Africa: meta-regression analyses.* Sci Rep, 2022. **12**(1): p. 20637.
43. Mahmud, S., et al., *The status of hepatitis C virus infection among people who inject drugs in the Middle East and North Africa.* Addiction, 2020. **115**(7): p. 1244-1262.
44. Mohamoud, Y.A., F.D. Miller, and L.J. Abu-Raddad, *Potential for human immunodeficiency virus parenteral transmission in the Middle East and North Africa: an analysis using hepatitis C virus as a proxy biomarker.* World J Gastroenterol, 2014. **20**(36): p. 12734-52.
45. Mahmud, S., et al., *Key associations for hepatitis C virus genotypes in the Middle East and North Africa.* J Med Virol, 2020. **92**(3): p. 386-393.
46. Chemaitelly, H., et al., *Who to Test for Hepatitis C Virus in the Middle East and North Africa?: Pooled Analyses of 2,500 Prevalence Measures, Including 49 Million Tests.* Hepatol Commun, 2019. **3**(3): p. 325-339.
47. Mumtaz, G.R., et al., *Status of the HIV epidemic in key populations in the Middle East and north Africa: knowns and unknowns.* Lancet HIV, 2022. **9**(7): p. e506-e516.
48. Benova, L., S.F. Awad, and L.J. Abu-Raddad, *Estimate of vertical transmission of Hepatitis C virus in Pakistan in 2007 and 2012 birth cohorts.* J Viral Hepat, 2017. **24**(12): p. 1177-1183.
49. Benova, L., et al., *Estimation of hepatitis C virus infections resulting from vertical transmission in Egypt.* Hepatology, 2015. **61**(3): p. 834-42.
50. Benova, L., et al., *Vertical transmission of hepatitis C virus: systematic review and meta-analysis.* Clin Infect Dis, 2014. **59**(6): p. 765-73.

For peer review only

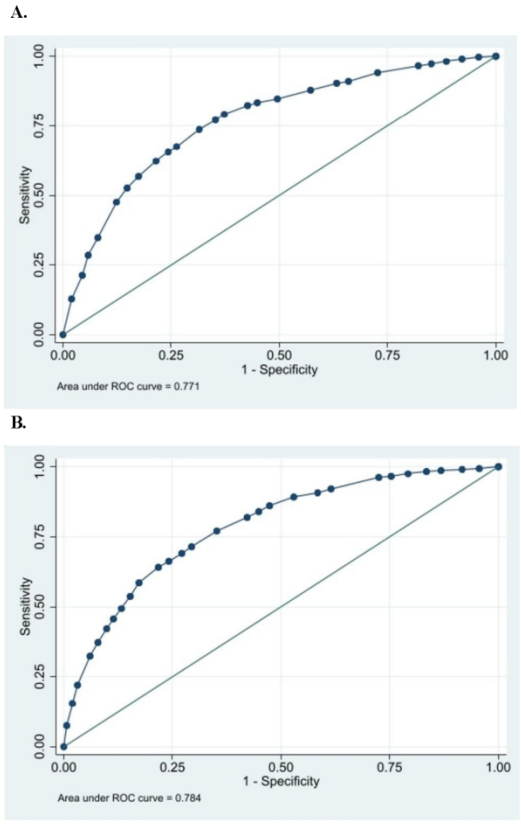
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Abbreviations: EDHS, Egypt Demographic and Health Survey; ROC, receiver operating characteristics.
The two scores exhibited comparable structures and diagnostic performances, with minor differences attributed to sampling variation of the same population across two distinct rounds of the EDHS surveys. Details on the derivation of these scores are provided in the Methods and Results sections.

Mathematical formula of the derived Egypt Hepatitis C Risk Score. A) Egypt Hepatitis C Risk Score using the EDHS 2008. B) Egypt Hepatitis C Risk Score using the EDHS 2014.

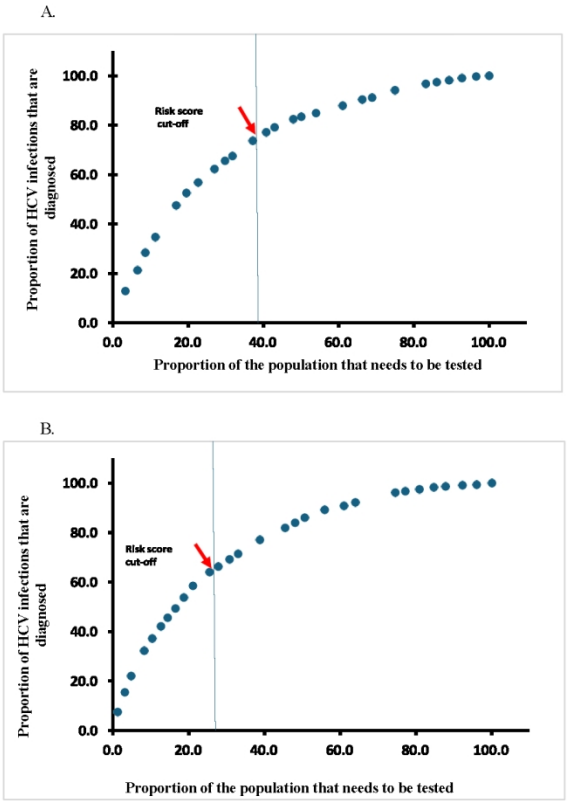
215x279mm (231 x 223 DPI)



Abbreviations: EDHS, Egypt Demographic and Health Survey; ROC, receiver operating characteristics.

Diagnostic performance of the Egypt Hepatitis C Risk Score using the area under the (ROC) curve. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

215x279mm (204 x 204 DPI)



Proportion of HCV infections in the population that are diagnosed as a function of the proportion of the population that needs to be tested to identify these infections. The figure shows the effect of prioritization of testing for those with higher to lower risk score. A) Using the EDHS 2008 data. B) Using the EDHS 2014 data.

215x279mm (200 x 200 DPI)

Table S1 Characteristics of individuals tested for HCV antibodies in the EDHS 2008 and 2014

Characteristics	EDHS 2008			EDHS 2014		
	Total tested (%)	HCV antibody positive (proportion %)	p-value	Total tested (%)	HCV antibody positive (proportion %)	p-value
No	11,126	1,571		26,047	1,456	
Sex						
Female	6,052 (54.4)	711 (11.8)	<0.001	13,707 (52.6)	660 (4.8)	<0.001
Male	5,074 (45.6)	860 (17.0)		12,340 (47.4)	796 (6.4)	
Age group (years)						
1-4	-	-	-	3,282 (12.6)	10 (0.3)	<0.001
5-9	-	-	-	3,601 (13.8)	10 (0.3)	
10-14	-	-	-	3,161 (12.1)	23 (0.7)	
15-19	2,000 (18.0)	82 (4.1)	<0.001	2,568 (9.9)	30 (1.2)	
20-24	1,837 (16.5)	91 (5.0)		1,976 (7.6)	54 (2.7)	
25-29	1,520 (13.7)	92 (6.1)		2,358 (9.0)	88 (2.7)	
30-34	1,244 (11.2)	133 (10.7)		2,076 (8.0)	114 (5.5)	
35-39	1,141 (10.3)	147 (12.9)		1,853 (7.1)	130 (7.0)	
40-44	1,069 (9.6)	238 (22.3)		1,468 (5.6)	146 (10.0)	
45-49	939 (8.4)	275 (29.3)		1,380 (5.3)	208 (15.1)	
50-54	728 (6.5)	272 (37.4)		1,334 (5.1)	337 (25.3)	
55-59	648 (5.8)	241 (37.2)		990 (3.8)	306 (30.9)	
Type of place of residence						
Urban	4,448 (40.0)	442 (9.9)	<0.001	11,955 (45.9)	546 (4.6)	<0.001
Rural	6,678 (60.0)	1,129 (16.9)		14,092 (54.1)	910 (6.5)	

Abbreviations: EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table S2 STROBE checklist for cross-sectional studies

	Item No	Recommendation	Main Text Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Abstract **page 2**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract **page 2**
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction **page 4**
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction **page 5**
Methods			
Study design	4	Present key elements of study design early in the paper	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') **pages 5-8**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods ('Egypt Demographic and Health Surveys') **pages 5-6**
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	Methods ('Egypt Demographic and Health Surveys') **pages 5-6**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') **pages 5-8** &Table S1 **page 1**
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') **pages 5-8** & Table 1& Table 3 & Tables S2-S3 **pages 3-4** & Figures 1-2
Bias	9	Describe any efforts to address potential sources of bias	Methods ('Risk score derivation'& 'Performance and validation of the risk score') **pages 6-8** & Table 2 & Figures 1-2
Study size	10	Explain how the study size was arrived at	Not applicable, see Methods ('Egypt Demographic and Health Surveys') **pages 5-6**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation'& 'Performance and validation of the risk score') **pages 5-8** & Tables 1-3 & Table S3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Methods ('Risk score derivation'& 'Performance and validation of the risk score') **pages 5-8**
		(*b*) Describe any methods used to examine subgroups and interactions	Not applicable, see Methods ('Risk score derivation'& 'Performance and validation of the risk score') **pages 6-8**
		(*c*) Explain how missing data were addressed	Not applicable, see Methods ('Egypt Demographic and Health Surveys', 'Risk score derivation') **pages 5-8**
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	Not applicable
		(*e*) Describe any sensitivity analyses	Not applicable
Results			
Participants	13	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	Table 1 & Tables S1 & Table S3
Descriptive data	14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 1 & Table S1 & Table S3 & Figure S1
		(b) Indicate number of participants with missing data for each variable of interest	Not applicable, see Methods ('Egypt Demographic and Health Surveys') **pages 5-6**
Outcome data	15	Report numbers of outcome events or summary measures	Results **pages 8-11**, Figures 1-2, Figure S1, Tables 1-3 & Table S1
Main results	16	(*a*) Give unadjusted estimates and, if applicable,	Table 1 & Table S3

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	Tables 1-3 & Table S1 & Table S3
		© If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Tables 2-3 & Table S3 & Figure S1
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Results **pages 8-11** & Table 2 & Figure 2
Discussion			
Key results	18	Summarise key results with reference to study objectives	Discussion **pages 11-12**, paragraphs 1-3
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion, paragraph 9 **page 14**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion, paragraphs 4-6 **pages 12-13**
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussion, paragraphs 7-8 **pages 13-14**
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Acknowledgements **page 16**

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table S3 Results of univariable and multivariable regression analyses to derive the Egypt Hepatitis C Risk Score using data from EDHS 2008 and EDHS 2014

Characteristics	EDHS 2008						EDHS 2014					
	OR* (95% CI)	p-value	aOR*† (95% CI)	p-value	β‡	Risk score§	OR* (95% CI)	p-value	aOR*† (95% CI)	p-value	β‡	Risk score§
Sex												
Female	1.00		1.00		Ref	0	1.00		1.00		Ref	0
Male	1.52 (1.35-1.70)	<0.001	1.52 (1.34-1.73)	<0.001	0.42	4	1.60 (1.41-1.83)	<0.001	1.62 (1.40-1.87)	<0.001	0.48	5
Age group (years)												
15-19	1.00		1.00			0	1.00		1.00			
20-24	1.19 (0.86-1.64)	0.282	1.23 (0.89-1.69)	0.213	0.20	2	3.18 (1.82-5.58)	<0.001	3.30 (1.88-5.81)	<0.001	1.19	12
25-29	1.52 (1.09-2.11)	0.014	1.60 (1.15-2.23)	0.005	0.47	5	4.38 (2.54-7.55)	<0.001	4.50 (2.60-7.79)	<0.001	1.51	15
30-34	3.09 (2.27-4.21)	<0.001	3.21 (2.35-4.39)	<0.001	1.17	12	7.33 (4.29-12.49)	<0.001	7.41 (4.35-12.65)	<0.001	2.00	20
35-39	3.69 (2.70-5.06)	<0.001	3.89 (2.84-5.34)	<0.001	1.36	14	8.56 (5.04-14.52)	<0.001	8.74 (5.13-14.88)	<0.001	2.17	22
40-44	6.91 (5.15-9.26)	<0.001	7.36 (5.47-9.90)	<0.001	1.99	20	12.54 (7.51-20.95)	<0.001	13.03 (7.79-21.81)	<0.001	2.57	26
45-49	9.30 (6.95-12.27)	<0.001	10.34 (7.71-13.85)	<0.001	2.34	23	18.64 (11.33-30.66)	<0.001	19.23 (11.66-31.69)	<0.001	2.96	30
50-54	14.36 (10.78-19.13)	<0.001	16.43 (12.29-21.96)	<0.001	2.80	28	37.25 (22.75-60.98)	<0.001	41.11 (25.05-67.46)	<0.001	3.71	37
55-59	15.09 (11.13-20.45)	<0.001	17.05 (12.50-23.26)	<0.001	2.84	28	49.26 (30.06-80.72)	<0.001	55.31 (33.59-91.06)	<0.001	4.01	40
Type of place of residence												
Urban	1.00		1.00		Ref	0	1.00		1.00		Ref	0
Rural	1.91 (1.66-2.19)	<0.001	2.34 (2.0-2.7)	<0.001	0.85	9	1.73 (1.46-2.1)	<0.001	2.15 (1.78-2.59)	<0.001	0.76	8

Abbreviations: aOR, Adjusted odds ratio; CI, Confidence interval; EDHS, Egypt Demographic and Health Survey; HCV, Hepatitis C virus; OR, Odds ratio; Ref, Reference category.
*The analysis applied the EDHS sampling weights.
†The odds ratio was adjusted for sex, age, and type of place of residence.
‡ β-coefficients were based on the multivariable regression analysis.
§The risk score was calculated by multiplying the β coefficient by 10 and then rounding the result to the nearest integer.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml