

## Refereed paper

# The comparison of cardiovascular risk scores using two methods of substituting missing risk factor data in patient medical records

Andrew RH Dalton MSc

PhD Student Public Health, Department of Primary Care and Public Health, Imperial College Faculty of Medicine, London, UK and Department of Public Health, NHS Ealing, Southall, UK

Alex Bottle PhD

Lecturer in Medical Statistics

Michael Soljak PhD

Research Associate Public Health

Department of Primary Care and Public Health, Imperial College Faculty of Medicine, London, UK

Cyprian Okoro

Consultant in Public Health, Department of Primary Care and Public Health, Imperial College Faculty of Medicine, London, UK and Department of Public Health, NHS Ealing, Southall, UK

Azeem Majeed PhD

Professor of Primary Care

Christopher Millett PhD

Senior Lecturer in Public Health

Department of Primary Care and Public Health, Imperial College Faculty of Medicine, London, UK

## ABSTRACT

**Background** Targeted screening for cardiovascular disease (CVD) can be carried out using existing data from patient medical records. However, electronic medical records in UK general practice contain missing risk factor data for which values must be estimated to produce risk scores.

**Objective** To compare two methods of substituting missing risk factor data; multiple imputation and the use of default National Health Survey values.

**Methods** We took patient-level data from patients in 70 general practices in Ealing, North West London. We substituted missing risk factor data using the two methods, applied two risk scores (QRISK2 and JBS2) to the data and assessed differences between methods.

**Results** Using multiple imputation, mean CVD risk scores were similar to those using default

national survey values, a simple method of imputation. There were fewer patients designated as high risk (>20%) using multiple imputation, although differences were again small (10.3% compared with 11.7%; 3.0% compared with 3.4% in women). Agreement in high-risk classification between methods was high (Kappa = 0.91 in men; 0.90 in women).

**Conclusions** A simple method of substituting missing risk factor data can produce reliable estimates of CVD risk scores. Targeted screening for high CVD risk, using pre-existing electronic medical record data, does not require multiple imputation methods in risk estimation.

**Keywords:** cardiovascular disease, electronic health records, health inequalities, primary prevention

## Introduction

Considerable resources have been devoted to reducing the burden of cardiovascular disease (CVD) in the UK over the past decade.<sup>1</sup> In England, this includes the National Health Service (NHS) Health Check programme, which began in 2008.<sup>2</sup> Despite this national commitment to a 'universal' programme, evidence is growing that a targeted approach to prevention might be an effective and cost-effective alternative.<sup>3,4</sup> Risk stratification involves deriving a CVD risk score using pre-existing data from patient medical records. A frequent problem, however, are missing data. Despite improvements in recording, there is still incompleteness.<sup>5,6</sup> When data are incomplete, default values must be entered to produce the risk score.

We compare two methods of data imputation; first, risk factor estimates derived from national survey data. This method is computationally simple and is used in bespoke CVD prevention software,<sup>7</sup> however, it has a limited evidence base and could create inaccuracies if local and national risk factor profiles differ. A second option is to generate estimates from local data using multiple imputation.<sup>8</sup> Multiple imputation is most effective when missing data are unrelated to observation characteristics [missing completely at random (MCAR)], however, it can also be effective when the pattern of missing data is dependent on recorded patient characteristics [missing at random (MAR)].<sup>9</sup> Multiple imputation is more computationally difficult, however, and is likely to be sensitive to patient differences and maintain the local risk profile within the imputed data. We compare imputation methods in both JBS2<sup>10</sup> and QRISK2<sup>11</sup> risk scores, analysing data from a deprived, ethnically diverse population. We further compare differences in risk stratification between risk scores concentrating on the practice level workload.

## Methods

### Data source

We obtained patient-level data from electronic medical records in 70 of the 85 general practices in Ealing, North West London. Data were extracted as part of the NHS Health Check programme, using Oberoi primary prevention software. We used data from the baseline of the programme,<sup>5</sup> extracted between December 2008 and December 2009. Briefly, data consist of demographic, anthropometric and clinical data relating to cardiovascular risk for patients registered in general practice, aged 35–74 years without diagnosed coronary

heart disease or stroke; Table 1 lists the variables extracted. We did not have data recorded for rheumatoid arthritis. Information on area deprivation is required for the QRISK2 risk engine, Townsend scores were linked to patient data using postcode of residence. We also linked 2007 Index of Multiple Deprivation (IMD) scores to data for the analyses. Data represent a patient's latest record at the time of extraction. We discounted data older than 15 years, because complete data were more important than timely data. The majority of data records were, however, from the previous five years.

The data contained missing values for a number of variables (Table 1). For both methods of imputation, we entered patients with missing ethnicity records as a subgroup in the analysis; for deprivation we replaced the missing Townsend score with the primary care trust (PCT) median. For rheumatoid arthritis, we assumed the condition to be absent if missing, in line with guidance for the QRISK2 score,<sup>11</sup> as were chronic kidney disease (CKD) and a family history of CVD.

We used two methods to estimate values for missing blood pressure, total cholesterol, high-density lipoprotein (HDL), body mass index (BMI) and smoking status. First, we used multiple imputation, using the *uvis* commands in Stata, generating 10 imputed datasets (*m*). We used multi-level linear imputation models (logistic regression for smoking status) to impute missing values. Level one of the model was the patient and level two was the general practice. Variables entered into models are shown in Table 1, with the addition of interaction terms between sex and age, and sex and ethnicity. We calculated the fraction of missing information ( $\gamma$ ) for each imputed variable.<sup>12</sup> For each variable imputed, given the magnitude of  $\gamma$ ,  $m = 10$  gave adequate power of imputation.<sup>12</sup> We dropped imputed values outside the range valid for the QRISK2 algorithm.

For the second method of substituting missing data we took mean risk factor values from the 2008 Health Survey for England (HSfE) in each year of age and sex group. For patients with missing risk factor data we entered the matching mean value from the HSfE based on the age and sex of the patient and assumed patients with missing smoking data to be non-smokers.

### Analysis

We compared CVD risk factors between ethnic groups using *t*-tests (Wilcoxon Mann–Whitney *U*-test for IMD), using the following five categories (missing, white, south Asian, black and other). We present age and deprivation breakdowns, plus age-standardised risk factor summaries for both men and women. We used the direct method of standardisation, using eight equal-age groups and the complete dataset as the

**Table 1** Variables included in QRISK2 and JBS2 score, and levels of missing data

	QRISK2	JBS2	% Missing
Age <sup>a</sup>	x	x	0.0
Sex <sup>a</sup>	x	x	0.0
Ethnicity <sup>a</sup>	x	c	34.3
Deprivation (Townsend score) <sup>a,b</sup>	x		11.9
Smoking	x	x	3.2
Hypertension <sup>a</sup>	x		0.0
Diabetes <sup>a</sup>	x	x	0.0
Blood pressure	x	x	11.1
Cholesterol/HDL	x	x	44.6
Body mass index	x		15.7
Family history of CVD <sup>a</sup>	x	c	28.6
Heart failure	x		18.5
Chronic kidney disease <sup>a</sup>	x		18.5
Rheumatoid arthritis	x		100
Atrial fibrillation	x		18.5

<sup>a</sup> Variables used in the multiple imputation of missing data. <sup>b</sup> Townsend score for QRISK2/IMD2007 for other analysis. <sup>c</sup> Not included within the algorithm, but adjustment factors are added.

standard. We applied the QRISK2<sup>13</sup> (after the 2010 update) and JBS2<sup>10</sup> risk scores to each imputed copy of the data set, calculating the mean score across imputations for each patient, for each method of imputation. We summarised levels of the two risk scores and differences overall in each sex/ethnicity group, and the proportion designated as at high risk ( $\geq 20\%$ ) – using direct age standardisation.

We compared the methods of data imputation. For the two risk scores, we calculated the mean risk score and the percentage of the population designated as high risk using the multiple imputation and health survey data. We assessed the agreement between the methods of the high-risk status using Cohen's Kappa, quoting 95% confidence intervals (CI).<sup>14</sup> Outside the population characteristics, we present the results for those patients with incomplete ( $n = 63\,607$ ) data recording to fully establish the impact of data imputation. All analyses were carried out using Stata 11.1. We obtained ethical approval for the use of anonymised patient-level data from the London Research Ethics Committee.

## Results

Table 2 shows the characteristics of the study population; the white population was more affluent than other ethnicities (Wilcoxon Mann–Whitney  $U$ -test  $P < 0.001$  compared with all other ethnic groups), smoking was higher in women ( $t$ -test  $P < 0.001$  compared with all ethnic groups). The south Asian and black populations have higher levels of diabetes ( $t$ -test  $P < 0.001$  compared with all other ethnic groups) and in the female population, black patients have the highest levels of obesity ( $t$ -test  $P < 0.001$  compared with all other ethnic groups).

In the entire population ( $n = 127\,724$ ), using multiple imputation data, the mean QRISK2 score [8.4 (95% CI = 8.4–8.4)] was significantly lower than JBS2 [11.1 (11.1–11.2)] with fewer patients designated as being at high risk [ $n = 15\,258$ ; 11.9% (11.8–12.1)] and  $n = 21\,377$ ; 16.7% (16.5–16.9), respectively] (Table A1). The QRISK2 score was lower in both sexes, and over most ethnic groups, with the largest difference in men of south Asian origin. Patients not of white, south

Asian or black ethnicity had a higher QRISK2 score than JBS2 in women.

In the population with missing risk factor data ( $n = 63\,607$ ), the multiple imputation method produces similar estimates of risk using both scores, but lower proportions of the population at high risk compared with using HSfE estimates. Differences were, however, relatively small. Using QRISK2 (Table 3), there were 10.3% at high risk using multiple imputation compared with 11.7% using the HSfE estimates in men; for women this was 3.0% compared with 3.4%, respectively (JBS2 risk score is shown in Table A2).

There were differences in mean risk scores and proportion at high risk using the two methods by sex and ethnic group. White men and women, south Asian men and black women were more likely to be designated as at high risk using National Health Survey data but these differences were small. In designating patients to be at high risk, the two methods showed a strong agreement with high kappa values across all sex and ethnic groups. In black men there was a significantly higher agreement between methods to other ethnic groups.

## Discussion

### Principal findings

We compared the cardiovascular risk scores generated for patients using two methods to substitute missing risk factor data. One approach was methodologically and computationally simple, using National Health Survey data, whilst the second was more complex using multiple imputation. Using multiple imputation the risk score estimates were similar and there was a small but significantly lower prevalence of high-risk status. Ethnic differences in risk scores were seen between the methods, although these were again only small. Multiple imputation is regarded as an accurate method of dealing with missing data and is robust to large amounts of missing data.<sup>8</sup> It is, however, a more time-consuming and complicated process than the use of default values from survey data, and is less transparent to non-statistically trained people. These small differences are unlikely to be clinically significant.

### What is already known on this topic?

Evidence has grown supporting targeted screening for CVD as both an effective and cost-effective approach. Interrogating general practice data, Marshall *et al* found that a large proportion of the patients eligible for primary prevention therapy lie in those with the

highest estimated risk.<sup>15</sup> Recent evidence suggests that pre-stratification of patient in the Health Check programme may alleviate the need for complete *de novo* risk factor recording.<sup>16</sup>

Modelling shows targeted screening as more cost-effective than a universal approach,<sup>3,4</sup> with limited value in recording new risk factor data in the whole population.<sup>17</sup> The call for targeted screening is not a new one and there is growing evidence to support this approach,<sup>18</sup> a recent case-control study has shown the strengths of using existing administrative data in the USA to predict stroke.<sup>19</sup> CVD risk factor recording is not complete in patient medical records,<sup>5</sup> but partial data maintains accurate risk prediction.<sup>20</sup>

### Limitations of the methods

Our data were extracted from the electronic medical records of a large number of patients. They cover an ethnically and socio-economically diverse population, different from those previously used to compare CVD risk scores in the UK,<sup>21</sup> and cover a large proportion of the registered patients in one English PCT. We used the recently updated version of the QRISK2 algorithm, and the modified Framingham (JBS2) risk score recommended as an alternative to QRISK2 in UK national clinical guidance.<sup>22</sup> Our data are timely and accurately represent the level of cardiovascular risk in a deprived, ethnically diverse population, although may not represent risk in other settings.

We cannot compare the predictive accuracy of the risk score as we were unable to link data to cardiovascular outcomes. The QRISK2 score may underestimate risk, especially in minority ethnic groups,<sup>11</sup> due to the small ethnic minority populations in the derivation data set. QRISK2 and its predecessor QRISK do, however, predict risk more accurately in the UK than Framingham risk scores.<sup>21</sup> Aside from missing data imputed in the study, our data set contained no data on rheumatoid arthritis and had missing data for atrial fibrillation, CKD and heart failure, all of which are variables in the QRISK2 score. These are not core components of the algorithm and the risk score allows them to be treated as absent if missing.<sup>11</sup> There were further missing data for the Townsend deprivation score which were replaced with the median score from the PCT. Neither method of data imputation will fully account for patient-level differences, it is likely those with missing data will differ from the remaining population. Although we use national default risk factor values as an exemplar, with primary care information systems in England, general practice mean values may prove easier to produce, and likely to be more reliable.

## Call for further research

Our data add to work showing the effectiveness of targeted screening by demonstrating that simple data methods can be used. In a UK setting, formal modelling to compare alternate CVD prevention strategies, such as a targeted one, with the NHS Health Check strategy is vital. Our dataset contains no data on CVD endpoint, including CVD mortality. Analyses linking methods of risk stratification and the management of missing risk factor data to predicted outcome will further show the value of such methods.

## Implications of findings for practice

CVD risk scores currently lie at the heart of one of England's foremost public health initiatives, the NHS Health Check programme; they therefore have significant implications for patient care. In the UK, there is currently a large national expenditure on a primary prevention programme for CVD with spending estimates between £180 million and £240 million per year.<sup>23</sup> Early evaluation of the programme has indicated that uptake and subsequent referrals are more limited than projected.<sup>24</sup> This will limit overall effectiveness.<sup>25</sup> If restricted patient involvement persists, especially given the need for greater efficiency in health spending, policy alternatives must be considered. Evidence from previous studies, outlined above, indicates that target CVD prevention is a viable alternative. This, however, was never considered when planning the NHS Health Check programme.<sup>23</sup>

CVD risk scores were designed as clinical tools, however, they are also effective at a population level,<sup>26</sup> and are the most efficient method of risk stratification.<sup>20</sup> Our findings demonstrate that a targeted screening method using electronic medical records in English primary care is easily implemented, and open to all providers, regardless of experience in informatics. Even if fully targeted screening is not adopted, the use of simple data imputation and stratification may create efficiency savings within a universal programme. Evidence, for example, suggests that pre-stratification may alleviate the need for cholesterol testing in all patients eligible for a Health Check.<sup>16</sup> The NHS Health Checks is, internationally, the first systematic, population-wide CVD prevention programme. Other countries with strong primary care structures and well-developed informatics may also consider the methods of targeting outlined here.

## Conclusion

Using existing medical record data for targeting screening, missing data are inevitable. A simple method of substituting missing data may be as effective in producing a CVD risk score as one using a complex method. A targeted approach to CVD prevention may be more cost-effective than a universal approach, and we demonstrated that it need not use complex methods to overcome missing data.

## ACKNOWLEDGEMENTS

We thank the 70 general practices providing data for the study.

## REFERENCES

- 1 Department of Health. *National Service Framework for Coronary Heart Disease*. Department of Health: London, 2000.
- 2 Department of Health. *Putting Prevention First, Vascular Checks; risk assessment and management*. Department of Health: London, 2008.
- 3 Chamnan P, Simmons RK, Khaw KT, Wareham NJ and Griffin SJ. Estimating the population impact of screening strategies for identifying and treating people at high risk of cardiovascular disease: modelling study. *BMJ* 2010;340:c1693.
- 4 Lawson KD, Fenwick EAL, Pell ACH and Pell JP. Comparison of mass and targeted screening strategies for cardiovascular risk: simulation of the effectiveness, cost-effectiveness and coverage using a cross-sectional survey of 3921 people. *Heart* 2010;96(3):208–12.
- 5 Dalton ARH, Bottle A, Okoro C, Majeed A and Millett C. Implementation of the NHS Health Checks programme: baseline assessment of risk factor recording in an urban culturally diverse setting. *Family Practice* 2011;28(1):34–40.
- 6 Majeed A, Car J and Sheikh A. Accuracy and completeness of electronic patient records in primary care. *Family Practice* 2008;25(4):213–14.
- 7 Oberoi. *Clinical Observations*. Oberoi Consulting: Derby, UK, 2007.
- 8 Janssen KJ, Donders AR, Harrell FE Jr *et al*. Missing covariate data in medical research: to impute is better than to ignore. *Journal of Clinical Epidemiology* 2010; 63(7):721–7.
- 9 Clark TG and Altman DG. Developing a prognostic model in the presence of missing data: an ovarian cancer case study. *Journal of Clinical Epidemiology* 2003;56(1): 28–37.
- 10 Joint British Societies (JBS2). Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. *Heart* 2005;91(Suppl. 5):v1–52.
- 11 Hippisley-Cox J, Coupland C, Vinogradova Y *et al*. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336(7659):1475–82.

- 12 Graham JW, Olchowski AE and Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science* 2007;8(3):206–13.
- 13 Dalton ARH, Soljak M, Samarasundera E, Millett C and Majeed A. Prevalence of cardiovascular disease risk amongst the population eligible for the NHS Health Check Programme. *European Journal of Cardiovascular Prevention and Rehabilitation* 2012. DOI: 10.1177/1741826711428797.
- 14 Cyr L and Francis K. Measures of clinical agreement for nominal and categorical data: the kappa coefficient. *Computers in Biology and Medicine* 1992;22(4):239–46.
- 15 Marshall T. The use of cardiovascular risk factor information in practice databases: making the best of patient data. *The British Journal of General Practice* 2006; 56(529):600–5.
- 16 Dalton ARH and Bull RJ. Risk stratification could reduce costs in primary prevention of cardiovascular disease. *BMJ* 2011;343:d4913.
- 17 Marshall T. Estimating the value of information in strategies for identifying patients at high risk of cardiovascular disease. *Informatics in Primary Care* 2006;14(2): 85–92.
- 18 Marshall T. Targeted case finding for cardiovascular prevention. *BMJ* 2010;340:c1376.
- 19 Ekundayo OJ, Vassar SD, Williams LS, Bravata DM and Cheng EM. Using Administrative databases to calculate Framingham scores within a large health care organization. *Stroke* 2011;42(7):1982–7.
- 20 Marshall T. Identification of patients for clinical risk assessment by prediction of cardiovascular risk using default risk factor values. *BMC Public Health* 2008;8:25.
- 21 Collins GS and Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ* 2009; 339:b2584.
- 22 National Institute of Clinical Excellence (NICE). *Cardiovascular Risk Assessment and the Modification of Blood Lipids for the Primary and Secondary Prevention of Cardiovascular Disease* (CG67). NICE: London, 2008.
- 23 Department of Health. *Economic Modelling for Vascular Checks*. Department of Health: London, 2008.
- 24 Dalton ARH, Bottle A, Okoro C, Majeed A and Millett C. Uptake of the NHS Health Checks programme in a deprived, culturally diverse setting: cross-sectional study. *Journal of Public Health* 2011;33(3):422–9.
- 25 Cooney MT, Dudina A, Whincup P et al. Re-evaluating the Rose approach: comparative benefits of the population and high-risk preventive strategies. *European Journal of Cardiovascular Prevention and Rehabilitation* 2009;16(5): 541–9.
- 26 Batsis JA and Lopez-Jimenez F. Cardiovascular risk assessment – from individual risk prediction to esti-

mation of global risk and change in risk in the population. *BMC Medicine* 2010;8(1):29.

#### CONFLICT OF INTERESTS

None declared.

#### FUNDING

This work was supported by the following; AD is funded by NHS Ealing. CM is funded by the Higher Education Funding Council for England and the National Institute for Health Research. MS was funded by the Economic & Social Research Council. AB and the Dr Foster Unit are largely funded by Dr Foster Intelligence, a joint venture with the Information Centre. The Department of Primary Care & Public Health at Imperial College is grateful for support from the National Institute for Health Research Biomedical Research Centre scheme, the National Institute for Health Research Collaboration for Leadership in Applied Health Research & Care scheme and the Imperial Centre for Patient Safety and Service Quality.

#### ETHICAL APPROVAL

London Research Ethics Committee.

#### CONTRIBUTORS AND SOURCES

AD, AB and CM conceived the study. AD conducted and AB supervised statistical analysis. AD wrote the first draft of the paper. CM, AB, MS, CO and AM helped interpret the data analysis and reviewed the manuscript critically for important intellectual content. AD will act as guarantor.

#### ADDRESS FOR CORRESPONDENCE

Andrew RH Dalton  
 Department of Primary Care and Public Health  
 Imperial College Faculty of Medicine  
 3rd Floor, Reynolds Building  
 St Dunstan's Rd  
 London W6 8RP  
 UK  
 Email: andrew.dalton@imperial.ac.uk

Accepted February 2012

## Appendix 1

**Table A1** Age-standardised risk scores using the multiple imputation data in the total population

	Missing	White	South Asian	Black	Other	Total
<b>Female</b>						
Mean risk score						
JBS2	6.6	7.6	8.1	7.6	7.7	7.4
QRISK	5.4	6.4	7.3	6.4	8.1	6.5
Difference	1.25	1.21	0.81	1.21	-0.43	0.95
(95% CI)	(1.22–1.29)	(1.17–1.27)	(0.76–0.87)	(1.12–1.29)	(-0.52– -0.33)	(0.92–0.97)
Proportion high risk						
JBS2	4.6	7.4	9.6	8.2	7.9	7.1
QRISK	3.8	6.7	9.8	6.9	11.7	7.1
Difference	0.84	0.65	-0.2	1.38	-3.85	0.02
(95% CI)	(0.59–1.1)	(0.34–0.96)	(-0.56–0.17)	(0.78–1.97)	(-4.4– -3.29)	(-0.14–0.17)
<b>Male</b>						
Mean risk score						
JBS2	13.1	14.2	20.4	13.8	14.4	15.1
QRISK	8.9	10.5	11.7	9.5	13.5	10.5
Difference	4.11	3.66	8.71	4.35	0.93	4.63
(95% CI)	(4.05–4.17)	(3.59–3.74)	(8.59–8.82)	(4.21–4.50)	(0.78–1.08)	(4.58–4.67)
Proportion high risk						
JBS2	21.0	24.4	40.6	23.9	25.6	26.7
QRISK	12.4	17.1	21.4	15.2	24.7	17.2
Difference	8.66	7.32	19.2	8.68	0.91	9.6
(95% CI)	(8.23–9.08)	(6.83–7.81)	(18.5–19.8)	(7.67–9.68)	(0.12–1.7)	(9.3–9.8)

**Table A2** Age-standardised JBS2 estimates using multiple imputation and HSfE data to substitute missing data in patients with missing data

	Missing	White	South Asian	Black	Other	Total
Risk score						
Male						
HSE	12.8 (12.7–12.9)	13.8 (13.7–14.0)	19.2 (18.9–19.5)	13.7 (13.4–14.0)	13.8 (13.5–14.0)	14.0 (13.9–14.0)
IMP	12.6 (12.5–12.7)	13.3 (13.1–13.4)	18.6 (18.3–18.9)	12.7 (12.4–13.1)	13.4 (13.2–13.7)	13.6 (13.5–13.7)
Female						
HSE	6.4 (6.3–6.4)	7.0 (6.9–7.1)	6.9 (6.7–7.1)	6.6 (6.4–6.8)	7.0 (6.8–7.2)	6.7 (6.6–6.7)
IMP	6.2 (6.1–6.2)	6.7 (6.6–6.8)	6.8 (6.6–7.0)	6.2 (6.0–6.4)	6.7 (6.5–6.9)	6.4 (6.4–6.5)
High risk						
Male						
HSE	21.7 (21.2–22.2)	24.7 (23.9–25.6)	37.3 (36.2–38.4)	23.8 (21.8–26)	23.5 (22.1–25.1)	24.5 (24.1–24.9)
IMP	20.1 (19.5–20.7)	21.4 (20.5–22.3)	36.2 (35–37.4)	21.1 (19–23.4)	22.6 (21.1–24.2)	22.6 (22.1–23.0)
Kappa	0.89 (0.87–0.90)	0.89 (0.88–0.89)	0.90 (0.90–0.91)	0.90 (0.89–0.91)	0.90 (0.90–0.91)	0.89 (0.89–0.90)
Female						
HSE	3.4 (3.0–3.8)	5.2 (4.6–5.8)	6.5 (5.5–7.7)	5.0 (3.8–6.6)	5.5 (4.5–6.6)	4.5 (4.2–4.8)
IMP	2.8 (2.5–3.2)	4.2 (3.6–4.8)	5.2 (4.3–6.3)	3.6 (2.6–5.0)	4.6 (3.7–5.7)	3.6 (3.4–3.9)
Kappa	0.89 (0.88–0.90)	0.85 (0.84–0.86)	0.83 (0.82–0.83)	0.88 (0.88–0.89)	0.88 (0.88–0.89)	0.87 (0.86–0.88)