Similar 9-Year Mortality Risks and Reproducibility for the World Health Organization and American Diabetes Association Glucose Tolerance Categories

The Hoorn Study

FEMMIE DE VEGT, MSC
JACQUELINE M. DEKKER, PHD
COEN D.A. STEHOUWER, MD
GIEL NIJPELS, MD
LEX M. BOUTER, PHD
ROBERT J. HEINE, MD

OBJECTIVE — To compare the risks of all-cause and cardiovascular disease (CVD) mortality in the American Diabetes Association (ADA) and World Health Organization (WHO) glucose tolerance categories after 9 years of follow-up in the Hoorn Study and to study the test–retest reproducibility of those categories.

RESEARCH DESIGN AND METHODS — In this population-based cohort study of 2,468 elderly men and women, subjects were classified according to both the WHO and the ADA criteria. Causes of death were extracted from the medical records. Age- and sex-adjusted relative risks were estimated by Cox's proportional hazards model. Reproducibility of the diagnostic criteria was assessed in a sample of 1,109 subjects with duplicate oral glucose tolerance tests.

RESULTS — Subjects with known diabetes had a four to five times higher risk of all-cause and CVD mortality compared with normal subjects ($P < 0.05$). The relative risks of all-cause mortality were 1.67 (95% CI 1.09–2.57) and 1.56 (1.00–2.43) for newly diagnosed diabetic subjects according to the WHO and ADA criteria, respectively. The WHO and ADA criteria had similar levels of reproducibility. The overall κ was 0.59 (0.54–0.64) for WHO criteria and 0.61 (0.56–0.66) for ADA criteria. For the category of newly diagnosed diabetes according to WHO or ADA, the percentages of agreement for the second test compared with the first test were 77% (85/110) and 74% (74/100), respectively.

CONCLUSIONS — Both sets of diagnostic criteria identify criteria-specific diabetic subjects with an increased mortality risk compared with normal subjects, and the reproducibility of both criteria is similar.

Diabetes Care 23:40–44, 2000

The American Diabetes Association (ADA) recently introduced new diagnostic criteria for the diagnosis of diabetes (1). In contrast to the 1985 World Health Organization (WHO) criteria (2), the ADA criteria are based on fasting plasma glucose (FPG) levels only, and consequently an oral glucose tolerance test (OGTT) is no longer required. Furthermore, ADA set the FPG cutoff point for the diagnosis of diabetes at 7.0 mmol/l (1).

Several studies have already shown poor agreement between the ADA and the WHO criteria. In the Decode Study, in which data from eight European countries were analyzed, the national changes in prevalence of diabetes ranged from a reduction of 4% to an increase of 13% when applying the new ADA diagnostic criteria (3). In the Hoorn Study, both sets of criteria led to similar prevalence figures, but 40% of the subjects newly diagnosed with diabetes according to the WHO criteria had no diabetes according to the ADA criteria (4). With either set of criteria, a considerable number of individuals with an adverse cardiovascular risk profile will not be diagnosed (4). Other studies also showed a large variation in individual classification (5–10).

At the time that the ADA criteria were introduced, no data were available regarding the mortality risks for the individuals diagnosed with diabetes by the new criteria. Although diagnostic criteria for diabetes are based on the presence of microvascular complications, diabetic individuals also are at risk for cardiovascular complications and cardiovascular mortality (11,12). Therefore, prospective data are needed to assess which set of criteria best identifies individuals at high risk for mortality and cardiovascular disease (CVD).

In epidemiological studies, the diagnosis of diabetes is often based on one abnormal diagnostic test. Both the ADA and the WHO criteria emphasize that, for use in clinical practice, the diagnosis of diabetes should always be confirmed by repeating the test on another day (1,2).

In the present study, we compared the risk of all-cause and CVD mortality in the ADA and WHO glucose tolerance categories after 9 years of follow-up in the Hoorn Study. Furthermore, we analyzed the test–retest reproducibility of the ADA and the WHO criteria.

From the Institute for Research in Extramural Medicine (E.D.V., J.M.D., C.D.A.S., G.N., L.M.B., R.J.H.), Vrije Universiteit; and the Department of Internal Medicine (C.D.A.S., R.J.H.), University Hospital Vrije Universiteit, Amsterdam, the Netherlands.

Address correspondence and reprint requests to F. de Vegt, MSc, Institute for Research in Extramural Medicine, Vrije Universiteit, Van der Boechorststraat 7, 1081 BT Amsterdam, the Netherlands.

Received for publication 13 July 1999 and accepted in revised form 5 October 1999.

Abbreviations: ADA, American Diabetes Association; CVD, cardiovascular disease; FPG, fasting plasma glucose; ICD-9, *International Classification of Diseases, Ninth Revision*; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; OGTT, oral glucose tolerance test; 2-h PG, 2-h postload glucose; WHO, World Health Organization.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

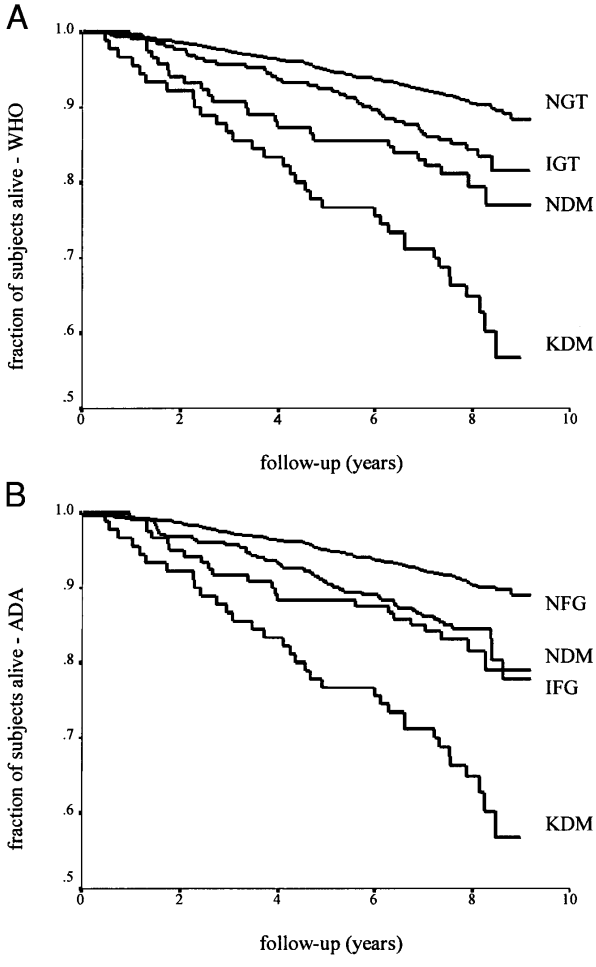

Figure 1—*Survival for the various glucose tolerance categories according to the WHO (*A*) and ADA (*B*) criteria. KDM, known diabetes; NDM, newly diagnosed diabetes; NFG, normal fasting glucose; NGT, normal glucose tolerance.*

RESEARCH DESIGN AND METHODS

Cohort population

The Hoorn Study, which began in 1989, is a population-based cohort study of glucose intolerance in a general elderly population. The study population has been described in detail (13). In short, a random sample of 3,553 men and women aged 50–75 years were invited to take part in the Hoorn Study; 2,540 subjects (71%) agreed. A total of 56 non-Caucasians were excluded from the study, which resulted in the final Hoorn Study cohort of 2,484 men and women. In the present study, 16 subjects were excluded because of missing glucose values, which resulted in a study population of 2,468 subjects. All subjects gave their written informed consent. The Ethics Committee of the University Hospital Vrije Universiteit approved the study.

Follow-up

The vital status of participants was provided by the Population Register of the city of Hoorn. Causes of death were extracted from the medical records of the general practitioners and the local hospital. The *International Classification of Diseases, Ninth Revision* (ICD-9) (14) was used to code the causes of death. CVD mortality was defined as ICD-9 codes 390–459 (diseases of the circulatory system) or 798 (sudden death, cause unknown) because sudden death in general is of CVD origin (15). A certified nosologist at the local hospital in Hoorn checked the ICD-9 codes assigned to the causes of death by one of the authors (F.d.V.) for a sample of 52 subjects; the agreement for the category of CVD mortality was high (94%). The vital status for the subjects who had moved out of Hoorn (n = 180) was checked in the population registers of the cities to which they had moved.

Glucose tolerance categories

At the baseline examinations (1989–1990), a blood sample was taken from all subjects after an overnight fast, after which a 75-g OGTT was administered. A glucose dehydrogenase method (Merck, Darmstadt, Germany) was used to determine FPG and 2-h postload glucose (2-h PG) levels. All subjects were classified in glucose tolerance categories according to the WHO and the ADA criteria (1,2). Subjects who were already registered as diabetic patients (i.e., subjects taking insulin or oral hypoglycemic agents or subjects treated with diet) were classified as having known diabetes.

In a subset of 1,109 subjects stratified for glucose tolerance status at the baseline examination, a second OGTT was performed after 2–6 weeks (16). These subjects were classified according to both the WHO and the ADA criteria (1,2) at both the first visit and at the second visit to determine the reproducibility of the ADA and the WHO criteria.

Statistical methods

To study possible differences in survival, Kaplan-Meier curves were made for the glucose tolerance categories according to both sets of criteria. Because age and sex are possible confounders in the association between glucose tolerance and mortality, age- and sex-adjusted relative risks (RRs) and 95% CIs were calculated by using Cox's proportional hazards model for the WHO and the ADA glucose tolerance categories separately. Subsequently, the same analyses were performed for glucose tolerance categories in which the WHO and the ADA criteria were combined (i.e., concordant and discordant glucose tolerance analyses were made).

In the subset of 1,109 subjects, the reproducibility of the WHO and the ADA criteria was examined by making cross-tables with the glucose tolerance categories of the first and second visits. The overall κ (which measures the agreement across the glucose tolerance categories at the first and second visits) and 95% CIs were calculated. A κ value >0.75 represents excellent agreement, values of <0.40 represent poor agreement, and values between 0.40 and 0.75 represent fair to good agreement (17).

All analyses were performed with the SPSS for Windows 6.1 software program (18). All P values were based on two-sided tests, and the cutoff for statistical significance was 0.05.

Table 1—*Age- and sex-adjusted RRs of all-cause and CVD mortality for glucose tolerance categories according to both ADA and WHO criteria*

Subjects	Cutoff values FPG/2-h PG (mmol/l)	*n*	Mortality %	All-cause	CVD
WHO criteria					
Nondiabetic	<7.8/and <7.8	2,008	9.4	1	1
IGT	<7.8/and 7.8–11.1	252	15.9	1.29 (0.91–1.82)	1.22 (0.73–2.05)
Newly diagnosed diabetes	≥7.8/or ≥11/1	118	20.3	1.67 (1.09–2.57)	1.52 (0.79–2.94)
Known diabetes	(by definition)	90	37.8	3.75 (2.60–5.42)	4.62 (2.77–7.71)
ADA criteria					
Nondiabetic	<6.1	1,973	9.3	1	1
IFG	6.1–7.0	285	16.5	1.53 (1.11–2.11)	1.24 (0.75–2.07)
Newly diagnosed diabetes	≥7.0	120	18.3	1.56 (1.00–2.43)	1.50 (0.78–2.89)
Known diabetes	(by definition)	90	37.8	3.83 (2.65–5.53)	4.62 (2.77–7.69)

Data are RRs (95% CIs).

RESULTS — The study population consisted of 1,137 men and 1,331 women with a mean age of 61.7 ± 7.3 years. During the 9 years of follow-up (until January 1999), 285 subjects died (11.6%). For 260 of those subjects (91.2%), the causes of death were retrieved; 131 of them (50.4%) had a CVD cause of death. Only 10 subjects (0.4%) were lost to follow-up, mainly because they had moved abroad.

The Kaplan-Meier curves show survival during the follow-up for glucose tolerance categories according to the WHO (Fig. 1*A*) and the ADA (Fig. 1*B*) criteria. The tendency is similar for both criteria, although the slopes for the ADA impaired fasting glucose (IFG) and newly diagnosed diabetes categories seem to be less distinct than the WHO impaired glucose tolerance (IGT) and newly diagnosed diabetes categories.

Table 1 shows age- and sex-adjusted RRs for all-cause and CVD mortality. Subjects with known diabetes ($n = 90$) had an approximately four times higher risk of all-cause mortality ($P < 0.05$). For subjects diagnosed with newly diagnosed diabetes according to the WHO criteria, the age- and sex-adjusted RR of all-cause mortality was 1.67 (95% CI 1.09–2.57), and this value was 1.56 (1.00–2.43) for newly diagnosed diabetic subjects according to the ADA criteria. The RR for IFG was almost as high (1.53 [1.11–2.11]). For CVD mortality, the age- and sex-adjusted RRs also increased for the categories of glucose intolerance, but only the RRs for known diabetes were statistically significant ($P < 0.05$).

When we combined both sets of criteria (Table 2), subjects with known diabetes had the highest age- and sex-adjusted risks of all-cause (3.95 [2.72–5.73]) and CVD (4.65 [2.77–7.80]) mortality. Discordant newly diagnosed diabetic subjects diagnosed by only one set of criteria had much lower RRs that were not statistically significant compared with concordant newly diagnosed diabetic subjects. However, the discordant newly diagnosed diabetes categories were very small and included only a few cases.

WHO and ADA criteria had a similar level of reproducibility (Tables 3 and 4). The overall κ values were 0.59 (0.54–0.64) for the WHO criteria and 0.61 (0.56–0.66) for the ADA criteria. The overall percentages of agreement were 81% ([706 + 112 + 85]/1109) for the WHO categories and 85% [(777 + 89 + 74)/1109] for the ADA categories. For the category of newly diagnosed diabetes according to the WHO or the ADA criteria, the agreement for the second test compared with the first test was 77 (85/110) and 74% (74/100), respectively, whereas for IGT (47%, 112/239) and IFG (51%, 89/173), the agreement was lower.

CONCLUSIONS — After 9 years of follow-up in the Hoorn Study, diabetic subjects newly diagnosed by either the ADA or the WHO criteria had an increased risk of all-cause mortality compared with normal subjects. The risk of CVD mortality was also increased, although this was not statistically significant.

The major reason to diagnose diabetes at an early stage is to prevent complications. Therefore, diagnostic criteria must identify subjects who are at high risk. The current WHO and ADA diagnostic criteria are based mainly on the association of hyperglycemia with microvascular complications (1,2). However, early treatment of hyperglycemia and associated CVD risk factors may also affect the severely enhanced risk of CVD and mortality.

In the present study, we did not find clear differences in mortality risks for subjects classified with IGT, IFG, or newly diagnosed diabetes according to either set of criteria, despite poor agreement between the ADA and the WHO criteria in the Hoorn Study as described before (4). Newly diagnosed diabetic subjects had an increased mortality risk of 1.56 with the ADA criteria (borderline significant) and 1.67 with the WHO criteria ($P < 0.05$).

Table 2—*Age- and sex-adjusted RRs of all-cause and CVD mortality for glucose tolerance categories concordant and discordant regarding ADA and WHO criteria*

Category	Cutoff values FPG/2-h PG (mmol/l)	*n*	Mortality %	All-cause	CVD
Normal (NFG/NGT)	<6.1/<7.8	1,791	8.8	1	1
IFG (ADA)/NGT (WHO)	6.1–7.0/<7.8	195	15.5	1.56 (1.06–2.31)	1.09 (0.56–2.12)
IGT (WHO)/NFG (ADA)	<6.1/7.8–11.1	163	14.7	1.25 (0.81–1.93)	1.05 (0.54–2.04)
IFG/IGT (ADA/WHO)	6.1–7.0/7.8–11.1	64	21.9	1.78 (1.03–3.09)	1.54 (0.67–3.54)
Newly diagnosed diabetes (ADA only)	7.0–7.8/<11.1	47	6.4	0.60 (0.19–1.87)	1.21 (0.38–3.85)
Newly diagnosed diabetes (WHO only)	<7.0/≥11.1	45	11.1	1.02 (0.42–2.48)	1.27 (0.40–4.04)
Newly diagnosed diabetes (ADA/WHO)	≥7.0 or ≥11.1	73	26.0	2.19 (1.35–3.53)	1.68 (0.77–3. 66)
Known diabetes	(by definition)	90	37.8	3.95 (2.72–5.73)	4.65 (2.77–7.80)

Data are RRs (95% CIs). NGT, normal glucose tolerance; NFG, normal fasting glucose.

Table 3—*Reproducibility of diagnostic criteria after 2–6 weeks in a sample of 1,109 subjects stratified according to glucose tolerance, according to WHO criteria*

WHO criteria

	First visit			
Second visit	NGT	IGT	Newly diagnosed diabetes	Total (%)
NGT	706	97	4	807 (72.8)
IGT	53	112	21	186 (16.8)
Newly diagnosed diabetes	1	30	85	116 (10.5)
Total (%)	760 (68.5)	239 (21.6)	110 (9.9)	1,109

NGT, normal glucose tolerance.

Table 4—*Reproducibility of diagnostic criteria after 2–6 weeks in a sample of 1,109 subjects stratified according to glucose tolerance, according to ADA criteria*

ADA criteria

	First visit			
Second visit	NFG	IFG	Newly diagnosed diabetes	Total (%)
NFG	777	67	4	848 (76.5)
IFG	56	89	22	167 (15.1)
Newly diagnosed diabetes	3	17	74	94 (8.5)
Total (%)	836 (75.4)	173 (15.6)	100 (9.0)	1,109

NFG, normal fasting glucose.

Until now, only two other large prospective studies compared the associations between mortality and diabetes according to the ADA and WHO criteria. In the Decode Study, which involved a combined analysis of 14 European cohort studies, no significant difference in mortality risk was evident for concordant and discordant diabetic men. In contrast, women with diabetes according to the 2-h PG levels had a mortality ratio in between those with fasting diabetes and those with concordant diabetes (19). In the Funagata Diabetes Study in Japan, the hazard ratios for all-cause and CVD mortality were higher for subjects diagnosed with diabetes according to the ADA criteria than for subjects diagnosed according to the WHO criteria. However, subjects with IGT had a higher risk of CVD mortality than subjects with IFG (20).

In the Hoorn Study (4), as in other studies (8,10), subjects with diabetes according to the ADA or the WHO criteria had an adverse CVD risk profile compared with subjects who had normal glucose values. The associated risks of mortality and CVD may be reduced by early treatment of CVD risk factors. If, in clinical practice, the fasting ADA criteria are used for diagnosis because of the convenience of requiring only a fasting blood sample, then individuals with isolated postload diabetes will not be diagnosed. Because the decision regarding whether to treat CVD risk factors is derived from risk stratification charts, the diagnosis of diabetes may be the decisive factor. Some risk stratification charts (21,22) present estimates of CHD risks separately for normal and diabetic individuals. Because individuals with diabetes have higher risk estimates, treatment of CVD risk factors is indicated earlier. Therefore, assessing postload glucose values in individuals at high risk for diabetes is advised, e.g., when the individual has a positive family history, is obese, or has an ethnic origin with a high prevalence of diabetes.

Reproducibility is an important issue in diagnosis. The present study is, as far as we know, the first study to compare directly the test–retest reproducibility of both the ADA and the WHO criteria in a large study population with duplicate OGTTs. Our sample of 1,109 subjects was stratified for glucose tolerance status at the first visit, which resulted in a higher prevalence of IGT, IFG, and diabetes compared with the general population. We observed a similar reproducibility of the WHO and the ADA criteria, as indicated by the κ values of 0.59 and 0.61, respectively, which represent fair to good agreement (17). The known high intra-individual variability of 2-h PG (16,23,24) did not result in a different reproducibility for the categorization based on the ADA and the WHO criteria.

The percentage of agreement for the diagnosis of diabetes was 77% when using the WHO criteria and 74% when using the ADA criteria. Yudkin et al. (25) showed a lower agreement of 64% for diabetes according to the WHO criteria in a population of 223 subjects also stratified for glucose tolerance status. Ko et al. (8) observed an agreement of 69% for subjects diagnosed with diabetes according to the ADA criteria. In the present study, the percentage of agreement for IGT and IFG was poor, as was observed in the studies mentioned above (8,25). This probably is due to the narrow ranges of FPG and 2-h PG defining these categories (25).

In conclusion, both sets of diagnostic criteria identify criteria-specific diabetic individuals with an increased mortality risk compared with nondiabetic individuals, and the reproducibility of both criteria is similar.

Acknowledgments — We thank the West Fries Gasthuis, the general practitioners, and the Population Register of the town of Hoorn for their cooperation with this study. We also thank Karien de Molennaar for her essential contributions to this article.

References

1. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus: Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 20 (Suppl. 1):1183–1197, 1997
2. World Health Organization: *Diabetes Mellitus: Report of a WHO Study Group*. Geneva, World Health Org., 1985 (Tech. Rep. Ser., no. 727)
3. Decode Study Group on behalf of the European Diabetes Epidemiology Study Group: Will new diagnostic criteria for diabetes mellitus change phenotype of patients with diabetes? Reanalysis of European epidemiological data. *BMJ* 317:371–375, 1998
4. de Vegt F, Dekker JM, Stehouwer CDA, Nijpels G, Bouter LM, Heine RJ: The 1997 American Diabetes Association criteria versus the 1985 World Health Organization criteria for the diagnosis of abnormal glucose tolerance: poor agreement in the Hoorn Study. *Diabetes Care* 21:1686–1690, 1998
5. Unwin N, Alberti KGMM, Bhopal R, Har-

land J, Watson W, White M: Comparison of the current WHO and new ADA criteria for the diagnosis of diabetes mellitus in three ethnic groups in the UK. *Diabet Med* 15: 554–557, 1998
6. Harris MI, Eastman RC, Cowie CC, Flegal KM, Eberhardt MS: Comparison of diabetes diagnostic categories in the U.S. population according to 1997 American Diabetes Association and 1980–1985 World Health Organization diagnostic criteria. *Diabetes Care* 20:1859–1862, 1997
7. Wahl PW, Savage PJ, Psaty BM, Orchard TJ, Robbins JA, Tracy RP: Diabetes in older adults: comparison of 1997 American Diabetes Association classification of diabetes mellitus with 1985 WHO classification. *Lancet* 352:1012–1015, 1998
8. Ko GTC, Chan JCN, Woo J, Cockram CS: Use of the 1997 American Diabetes Association diagnostic criteria for diabetes in a Hong Kong Chinese population. *Diabetes Care* 21:2094–2097, 1998
9. Gómez-Pérez FJ, Aquilar-Salinas CA, López-Alvarenga JC, Perez-Jauregui J, Guillen-Pineda LE, Rull JA: Lack of agreement between the World Health Organization category of impaired glucose tolerance and the American Diabetes Association category of impaired fasting glucose. *Diabetes Care* 21:1886–1888, 1998
10. Gimeno SGA, Ferreira SRG, Franco LJ, Iunes M: Comparison of glucose tolerance categories according to World Health Organization and American Diabetes Association diagnostic criteria in a population-based study in Brazil. *Diabetes Care* 21: 1889–1892, 1998
11. Adlerberth AM, Rosengren A, Wilhelmsen L: Diabetes and long-term risk of mortality from coronary and other causes in middle-aged Swedish men. *Diabetes Care* 21: 539–545, 1998
12. Jarrett RJ, McCartney P, Keen H: The Bedford Survey: ten year mortality rates in newly diagnosed diabetics, borderline diabetics and normoglycaemic controls and risk indices for coronary heart disease in borderline diabetics. *Diabetologia* 22:79–84, 1982
13. Mooy JM, Grootenhuis PA, de Vries H, Valkenburg HA, Bouter LM, Kostense PJ, Heine RJ: Prevalence and determinants of glucose intolerance in a Dutch Caucasian population: the Hoorn Study. *Diabetes Care* 18:1270–1273, 1995
14. World Health Organization: *International Classification of Diseases*. Vol. 1–2, 9th ed. Geneva, World Health Org., 1977
15. Kannel WB, Plehn JF, Cupples LA: Cardiac failure and sudden death in the Framingham Study. *Am Heart J* 115:869–875, 1988
16. Mooy JM, Grootenhuis PA, de Vries H, Kostense PJ, Popp-Snijders C, Bouter LM, Heine RJ: Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. *Diabetologia* 39: 298–305, 1996
17. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 33:159–174, 1977
18. Norusis MJ: *SPSS for Windows 6.1*. Chicago, SPSS, 1990
19. Balkau B for the Decode Study Group on behalf of the European Diabetes Epidemiology Study Group: New diagnostic criteria for diabetes and mortality in older adults. *Lancet* 353:68–69, 1999
20. Tominaga M, Eguchi H, Manaka H, Igarashi K, Kato T, Sekikawa A: Impaired glucose tolerance is a risk factor for cardiovascular disease, but not impaired fasting glucose: the Funagata Diabetes Study. *Diabetes Care* 22: 920–924, 1999
21. Wood D, de Backer G, Faergeman O, Graham I, Mancia G , Pyörälä K, with members of the Task Force: Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and Other Societies on Coronary Prevention. *Eur Heart J* 19:1434–1503, 1998
22. Yudkin JS, Charturvedi N: Developing risk stratification charts for diabetic and nondiabetic subjects. *Diabet Med* 16:219–227, 1999
23. Feskens EJM, Bowles CH, Kromhoud D: Intra- and interindividual variability of glucose tolerance in an elderly population. *J Clin Epidemiol* 44:947–953, 1991
24. Forrest RD, Jackson CA, Yudkin JS: The abbreviated glucose tolerance test in screening for diabetes: the Islington Diabetes Survey. *Diabet Med* 5:557–561, 1988
25. Yudkin JS, Forrest RD, Jackson CA: Misclassification of diabetic subjects may account for the increased vascular risk of impaired glucose tolerance: the Islington Diabetes Survey. *Diabetes Res Clin Pract* 13: 1–14, 1991