Cohort Profile: The Diabetes Study of Northern California (DISTANCE)—objectives and design of a survey follow-up study of social health disparities in a managed care population[†]

Howard H Moffet[1,*], **Nancy Adler**[2], **Dean Schillinger**[3], **Ameena T Ahmed**[1], **Barbara Laraia**[4], **Joe V Selby**[1], **Romain Neugebauer**[5], **Jennifer Y Liu**[1], **Melissa M Parker**[1], **Margaret Warton**[1], and **Andrew J Karter**[1,6]

1 *Kaiser Permanente – Division of Research, 2000 Broadway, Oakland, CA 94612, USA*

2 *Department of Psychiatry, Department of Pediatrics and Center for Health and Community, University of California, 3333 California Street, Suite 465, San Francisco, CA 94118, USA*

3 *University of California, San Francisco School of Medicine, Center for Vulnerable Populations, San Francisco General Hospital, San Francisco, CA 94110, USA*

4 *Department of Medicine and Center for Health and Community, University of California, 3333 California Street, Suite 465, San Francisco, CA 94118, USA*

5 *Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94704, USA*

6 *Department of Epidemiology, School of Public Health & Community Medicine, University of Washington, Seattle, WA 98195, USA*

How did the study come about?

One of the challenges of the national initiative, Healthy People 2010,[1] is to support interventions that will reduce social disparities in health. While social disparities such as differences in education, income, race or ethnicity may affect health, the mechanisms are poorly understood. If social disparities in health originate in childhood, are current social disparities in health modifiable and are they the responsibility of a medical provider or health plan? Nonetheless, modifiable factors may exist at the individual, neighbourhood or system level that mediate (explain) social disparities in health and that may be suitable targets for interventions aiming to reduce disparities. Our aim was to survey and prospectively follow a large, diverse and well-characterized population with diabetes and to collect data on risk factors which may affect diabetes health outcomes but which may differ substantively in prevalence or effect size across ethnic groups or educational levels.

The 2002 Institute of Medicine report, 'Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care',[2] detailed the socioeconomic fragmentation of health care quality and access, and its differential, negative impact on minorities. As access to health care is an important determinant of health outcomes and is in turn associated with ethnicity and socioeconomic position, it represents a potent source of confounding bias in population-based studies of social health disparities.

At the population level, identifying antecedents and determinants of social differences in disease progression remains a challenge. There are nationally representative surveys (e.g. National Health and Nutrition Examination Survey) that identify members of the general population with diabetes and that can estimate risk factors that are associated with diabetes complications, but these cohorts are often cross-sectional or panel studies and have limited ability to track continuous changes in individual-level health care, outcomes and risk factors over time. Longitudinal, population-based studies often have too few minority subjects with diabetes to reveal racial or ethnic differences.

The relatively uniform access to care in a managed care health plan such as Kaiser Permanente provides an advantageous setting in which to conduct a longitudinal study of social disparities in health and study findings may be compared with population-based studies, where health care access and quality may vary by socioeconomic position.

We established the Kaiser Permanente Northern California Diabetes Registry ('Registry') in 1993 using standardized criteria (Table 1) to identify and prospectively follow members with diabetes, to measure prevalence and incidence of diabetes and its co-morbidities, to understand factors associated with disease progression and complications, and to evaluate health care processes and outcomes. The Registry has an estimated sensitivity of 99% based on chart review validation (unpublished results). We conducted the first survey of the Registry in 1994–97 (Diabetes Registry Questionnaire) among all Registry members over 19 years of age. The primary goal of that survey was to capture individual-level information on the clinical characteristics of diabetes, age at diagnosis, ethnicity, education, health-related behaviours and diabetes family history. There were 77 726 respondents (83% response rate among eligible members) and that survey cohort has been the basis for numerous publications regarding the epidemiologic and health services aspects of diabetes.[3–9]

We previously reported findings regarding ethnic disparities in the incidence of myocardial infarction, stroke, congestive heart failure, end-stage renal disease and lower-extremity amputation among diabetic African American, Asian, Latino and Caucasian members of Kaiser Permanente Northern California ('Kaiser'),[7] a population with uniform access to care.[10] Socioeconomic disparities in diabetic complications based on educational attainment and income have been reported in other populations.[11,12]

The National Institutes of Health provided funding to the Kaiser Division of Research and the University of California, San Francisco School of Medicine, to conduct the Diabetes Study of Northern California (*DISTANCE*). This study was approved by their respective Institutional Review Boards.

What does it cover?

We developed and implemented DISTANCE as a survey follow-up cohort study[13] to assess a wide range of social and behavioural factors that we hypothesized to be potentially confounding, moderating or mediating factors associated with social disparities in diabetes-related outcomes. The DISTANCE Survey consisted of 184 questions and the domains included demographics, socioeconomics, clinical profile, health behaviours, treatment adherence, diabetes knowledge, psychosocial characteristics, patient–provider relationship and quality and access to care (Table 2). Every effort was made to create a survey with language, font and layout that would be accessible to non-English speakers and persons of low literacy or visual acuity.[14,15] The complete survey is available as Supplementary data and at http://distancesurvey.org.

Survey content

Our two primary exposures/attributes of interest were self-reported race/ethnicity and education (years of education and degrees earned). Other social factors surveyed included country of birth, acculturation and language fluency, subjective socioeconomic position[16,17] and several objective measures of socioeconomic position: educational attainment of self,[18,19] parents and spouse,[20] individual employment,[21,22] household income,[23] assets,[23,24] marital status and family size.

The survey included extensive questions on health behaviours and self-reported symptoms not readily available in administrative data: diet,[25] physical activity,[26] smoking,[21] alcohol consumption,[27] self-monitoring of blood glucose (SMBG),[25] oral health,[28,29] self-examination of feet,[25] medication adherence,[25,30,31] multivitamin use,[32] TV watching,[20] attitudes and beliefs about diabetes,[23] internal vs external locus of control,[33,34] social support,[23,35–37] health status,[38,39] stress,[40] depression,[22,41–43] sleep quality,[44–46] chronic pain,[38] personality traits,[47–49] health literacy,[50–52] diabetes knowledge,[53,54] medical and dental coverage, medical costs and ability to pay,[55] medical visit travel time,[43] language barriers, delayed treatment,[56] provider recommendations regarding aspirin[43] and SMBG,[43] foot exam by physician,[43] quality of care,[43] discrimination,[57,58] provider communication and interactions,[59,60] trust in provider[61,62] and self-reported height, weight,[43] symptoms, complications,[43] erectile dysfunction or female urinary incontinence.

Survey modes

The survey included four modes of administration: (i) a computer-assisted telephone interview (CATI) administered by a third party, (ii) a password-enabled, internet-accessible survey ('web survey') maintained on a secure server at the Kaiser Division of Research, (iii) a self-administered, written survey or (iv) a short version of the written survey. The content of each survey mode was identical except for slight adjustments in wording and logic patterns as needed and was estimated to take 45–60 min to complete; the short written version was abridged and contained 40 questions. The written and web surveys were in English only, but the CATI was available in English, Spanish, Cantonese, Mandarin and Tagalog using certified translations of an English script and was intended to maximize accessibility to those with language barriers or limited English literacy or fluency.

Other baseline data

We geocoded the home addresses of all respondents and non-respondents and linked their census block group and tract to 2000 Census data. These census characteristics were then factor-analysed, and we developed an eight-item 'deprivation index' as a way of summarizing the many indicators. The deprivation index has excellent internal reliability (Cronbach alpha = 0.93) with each of the eight items loading almost equally onto the underlying construct.

Baseline clinical characterizations were obtained from the extensive Kaiser administrative databases including smoking status, body mass index, current pharmacotherapy utilization and adherence,[63] laboratory findings, history of co-morbid events and procedures, use of emergency room, outpatient and inpatient health services, costs, outpatient and inpatient risk scores based on health care utilization and severity of disease[64] and end-stage renal disease registry (linked to the United States Renal Data System). These data are available for all subjects, both respondents and non-respondents, and allow for a substantive assessment of survey response bias.

Who is in the sample?

This study took place within the Kaiser Foundation Health Plan, the largest not-for-profit managed health care company in the United States.[65] Study subjects were members of Kaiser Permanente Northern California (KPNC) which provides comprehensive medical services to over 3.2 million members (as of January, 2005) in the San Francisco Bay and Sacramento metropolitan areas, or 25–30% of the region's population. KPNC members are predominantly employed or retired individuals and their families, and closely approximate the general population ethnically and socioeconomically except for the extreme tails of income distribution.[6,66,67] The Diabetes Registry consisted of 199 123 members as of January 1, 2005, and from this we selected an ethnically stratified, random sample of 40 735 health plan members, aged 30–75 at baseline, to receive the DISTANCE Survey: 6871 African American (17%), 11 197 Asian (27%), 4233 Caucasian (10%), 7018 Latino (17%) and 11 417 members of unknown ethnicity (28%). The DISTANCE Survey was in the field from May 5, 2005 until December 31, 2006.

The survey was completed by 20 188 persons (Table 3). The participation by ethnicity was 3420 African Americans (16.9%), 2312 Asian (11.4%), 4602 Caucasians (22.8%), 2404 Filipinos (11.9%), 3717 Latinos (18.4%), 2222 multi-racial (11.0%) and 1511 South Asian, Pacific Islander, Native America, Eskimo or other/unknown (7.5%). The distribution by mode was 10 429 CATI (51.7%), 4288 written survey (21.2%), 2393 short version (11.8%) and 3078 web (15.2%). Using an algorithm endorsed by the Council of American Survey Research Organizations (CASRO), if persons unable to be contacted had the same rate of eligibility as those contacted and were counted in the denominator, the survey response rate was 62%. Of the 20 188 respondents to this current survey, 4524 subjects (22%) had also responded to the prior Diabetes Registry Questionnaire (1994–97).

What has been measured?

The entire cohort has been characterized using demographic, clinical and behavioural data from administrative records and census data.

African American respondents were more likely to be female, to have had English as their preferred language, to have had elevated levels of low-density lipoprotein (LDL), uncontrolled blood pressure, were current smokers, with higher co-morbidity scores and living in a deprived neighbourhood.

Asian respondents were more likely to be male, over 60 years of age, to not have had English as their preferred language, to have had lower levels of LDL, better blood pressure control, lower missed appointment rate, better medication adherence, were non-smokers, not practicing SMBG, with lower co-morbidity scores and not living in an economically deprived or working class neighbourhood.

Caucasian respondents were more likely to complete the survey online, to have had lower mean A1C, practicing SMBG and not living in linguistically isolated neighbourhood.

Filipino respondents were more likely to participate via written survey, to have been older than 45 years of age, to have had lower levels of LDL and lower co-morbidity scores.

Latino respondents were more likely to participate via CATI, to have been younger than 45 years of age, to not have English as their preferred language, to have had higher mean A1C, higher rates of missed appointments, poorer medication adherence and living in a linguistically isolated or working class neighbourhood.

What outcomes will be measured during follow-up?

This cohort will be the basis for longitudinal evaluations of a wide range of clinical outcomes associated with diabetes and are powered to evaluate the ethnic and educational disparities in diabetes-related complication rates (e.g. myocardial infarction, stroke, heart failure, kidney failure and amputation) and mortality after a 2.5-year follow-up. Clinical and administrative follow-up data will be captured from these same sources as baseline, plus deaths from the state mortality files. Data from all sources combined may offer insights into how differences at the patient-level (e.g. differences in behaviours, type of therapy, intermediate health status and psychosocial factors, and competing demands), provider-level (e.g. trust in provider, racial concordance and patient-provider communication) and system-level (e.g. differential impact of cost-sharing) may lead to social disparities in diabetes health, end-stage diabetes complications and mortality.

What is the anticipated attrition?

One of the strengths of this diabetes registry is the very low turnover rate (~5% discontinue membership each year), affording minimal loss to follow-up. On average, the duration of Kaiser membership among our diabetes cohort members is 8.9 years (SD = 6.6).

What are the strengths and weaknesses?

A significant weakness is that written and web surveys were only offered in English. However, telephone interviewers made the initial efforts to contact subjects who had the opportunity to complete interviews in Spanish, Mandarin, Cantonese or Tagalog. Offering the survey by oral interview in multiple languages was intended to mitigate the language and/or literacy barriers, but we still observed lower participation rates by those with less education and modest response differences by race. On a more favourable note, 84% of non-respondents were identified by the health plan as having English as their preferred language, suggesting that language alone was not a major barrier to participation.

The response rate was, as expected, lower than our previous Diabetes Research Questionnaire (1994–97). A recent survey conducted in the Translating Research Into Action for Diabetes (TRIAD) study among managed care populations across the United States, including Kaiser, [43] had a response rate of ~50%. The Behavioural Risk Factor Surveillance System also experienced a decline in median response rates, from 63.2% in 1996 to 53.5% in 2001.[68] Privacy concerns and competition from telemarketing have probably played a role in declining response rates. However, if generally lower response rates have been observed in minority populations, the response rate for this survey may be considered favourable given that 77% of our sample were racial or ethnic minority health plan members.

One strength of this study is the relatively uniform access to care provided by membership in this integrated health plan. Thus, this study largely avoids disparities in access and quality, which are potent sources of confounding bias in population-based studies of social health disparities. Another strength is the quality and richness of the clinical follow-up data. Administrative and clinical data are available directly from the health plan and are considered to be reliable and complete. Each telephone interview was conducted by professional interviewers employed, trained and supervised by the Public Health Institute Survey Research Group in Sacramento, CA, which has its own internal quality controls. Written surveys (long- and short-forms) were coded and edited using an extensive data dictionary and detailed coding rules to standardize and clarify responses before being sent to the Data Entry department at the Kaiser Division of Research, which conducted double data entry. The survey data quality is limited by being self-reported.

The diversity, size and wealth of data included in the DISTANCE cohort make it suitable for the prospective study of a wide range of social disparities in the processes and outcomes of diabetes health care. Information about childhood socioeconomic position and educational attainment will facilitate a life course approach.[69,70]

Analysis of response bias

Response bias is a concern in any survey. Surveys are subject to response bias if participation is associated with the risk factors (exposures) and/or diseases (outcomes) under investigation. In this study population we have substantial data on non-respondents, including Kaiser administrative data, co-morbidity and census data on neighbourhood characteristics, allowing us to compare hypothesized relationships in respondents vs non-respondents and adjust for this potential response bias.[71] While participation was somewhat lower among minorities and those with less education, overall, respondents and non-respondents were quite similar.

We had pre-baseline reported race data for 29 319 (72%) members of the DISTANCE cohort and observed differences in response rates by race: African Americans 52%, Asians 46%, Caucasians 63%, Latinos 53%, unknown 45% ($P < 0.0001$). We also had educational attainment data for 4524 subjects (22%) who had completed the previous survey in 1994–97 and observed that subjects with a high school education or less participated at a lower rate than those with more than a high school education (54 vs 60%) ($P < 0.0001$).

A more complete view of response bias was obtained from health plan administrative data and census data, which was available for all members of the cohort. We found few notable differences between respondents and non-respondents in demographic, clinical, behavioural or neighbourhood characteristics. Compared with respondents, non-respondents had higher mean LDL levels (99.3 vs 94.3 mg/dl) and were less likely to practice SMBG (32 vs 46%). However, a portion of these unadjusted differences are likely due to age differences in respondents vs non-respondents.

We conducted an assessment of response bias by comparing differences in A1C across ethnic groups separately among respondents and non-respondents. We specified a logistic regression model of poor glycemic control which regressed the outcome, A1C >7%, on ethnicity, age and sex and an interaction term (ethnicity × survey response). Although respondents consistently had higher A1C than non-respondents, ethnic differences in the relationship between ethnicity and A1C did not differ between respondents and non-respondents ($P = 0.55$). We conducted a similar analysis with the subjects for whom we had educational attainment data from a previous survey and again found no substantive response bias ($P = 0.28$). While characteristics differ between respondents and non-respondents (response bias), the associations with outcomes of interest (slopes) are much less vulnerable to such bias.[72,73]

Statistical methods

Analyses of DISTANCE data will focus on associations rather than descriptive characterization. We will employ a modelling approach using inverse weighting estimation of marginal structural models[72] to investigate causal relationships in this study. These models will include weights to adjust for the non-proportionate sampling fractions (e.g. over-sampling of the minority ethnic groups) and response bias (e.g. giving greater weight to respondents who have characteristics more similar to the non-respondents). Thus, the proposed estimating approach will adjust for confounding and selection bias simultaneously.[73–77]

How can I collaborate? Where can I find out more?

We are not free to release participants' personal data under our promise to them regarding confidentiality. However, the DISTANCE steering committee is interested in collaborations

with external researchers. DISTANCE investigators are particularly interested in comparing social disparities observed in this insured population with uniform access to care to disparities observed in population-based samples, where quality and access to care vary widely by social strata. Requests for collaboration must be submitted to the director of the central coordinating centre (corresponding author). Before collaborations can be initiated, proposals require review and approval by the DISTANCE Publications and Presentations (P&P) Committee. This committee was formed to: (i) Ensure accurate, uniform, timely and high quality reporting of DISTANCE findings; (ii) Preserve the scientific integrity of the study and (iii) Safeguard the rights and confidentiality of participants.

Supplementary Data

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Healthy People 2010. U S Department of Health and Human Services, 2005. [Accessed August 2, 2007]. Available at: http://www.healthypeople.gov/

2. Smedley, BD.; Stith, AY.; Nelson, AR. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: Institute of Medicine, National Academy Press; 2002.

3. Karter AJ, Newman B, Rowell S, et al. Large-scale collection of family history data and recruitment of informative families for genetic analysis. J Reg Manag 1998;25:7–12.

4. Karter AJ, Ferrara A, Darbinian J, Ackerson LM, Selby JV. Self-monitoring of blood glucose: language and financial barriers in a managed care population with diabetes. Diabetes Care 2000;23:477–83. [PubMed: 10857938]

5. Karter AJ, Moffet HH, Liu J, et al. Achieving good glycemic control: initiation of new antihyperglycemic therapies in patients with type 2 diabetes from the Kaiser Permanente Northern California Diabetes Registry. Am J Manag Care 2005;11:262–70. [PubMed: 15839186]

6. Karter AJ, Ackerson LM, Darbinian JA, et al. Self-monitoring of blood glucose levels and glycemic control: the Northern California Kaiser Permanente Diabetes registry*. Am J Med 2001;111:1–9. [PubMed: 11448654]

7. Karter AJ, Ferrara A, Liu JY, Moffet HH, Ackerson LM, Selby JV. Ethnic disparities in diabetic complications in an insured population. JAMA 2002;287:2519–27. [PubMed: 12020332]

8. Karter AJ, Parker MM, Moffet HH, et al. Missed appointments and poor glycemic control: an opportunity to identify high-risk diabetic patients. Med Care 2004;42:110–15. [PubMed: 14734947]

9. Selby JV, Karter AJ, Ackerson LM, Ferrara A, Liu J. Developing a prediction rule from automated clinical databases to identify high-risk patients in a large population with diabetes. Diabetes Care 2001;24:1547–55. [PubMed: 11522697]

10. Martin TL, Selby JV, Zhang D. Physician and patient prevention practices in NIDDM in a large urban managed-care organization. Diabetes Care 1995;18:1124–32. [PubMed: 7587847]

11. Chaturvedi N, Jarrett J, Shipley MJ, Fuller JH. Socioeconomic gradient in morbidity and mortality in people with diabetes: cohort study findings from the Whitehall Study and the WHO Multinational Study of Vascular Disease in Diabetes. BMJ 1998;316:100–5. [PubMed: 9462313]

12. Booth GL, Hux JE. Relationship between avoidable hospitalizations for diabetes mellitus and income level. Arch Intern Med 2003;163:101–6. [PubMed: 12523923]

13. Kleinbaum, DG.; Kupper, LL.; Morgenstern, H. Epidemiologic Research: Principles and Quantitative Methods. New York: Van Norstrand Reinhold; 1982.

14. Dillman, DA. Mail and Internet Surveys: The Tailored Design Method. Vol. 2. New York: John Wiley; 2000.

15. Rudd RE, Comings JP, Hyde JN. Leave no one behind: improving health and risk communication through attention to literacy. J Health Commun 2003;8(Suppl 1):104–15. [PubMed: 14692575]

16. Singh-Manoux A, Adler NE, Marmot MG. Subjective social status: its determinants and its association with measures of ill-health in the Whitehall II study. Soc Sci Med 2003;56:1321–33. [PubMed: 12600368]

17. John, D.; Catherine, T. MacArthur Research Network on Socioeconomic Status and Health. The MacArthur Scale of Subjective Social Status. San Francisco: University of California; 1999 [Accessed July 12, 2007]. Available at: http://www.macses.ucsf.edu/Research/Psychosocial/notebook/subjective.html

18. Ross CE, Wu CL. Education, age, and the cumulative advantage in health. J Health Soc Behav 1996;37:104–20. [PubMed: 8820314]

19. Ross CE, Mirowsky J. Refining the association between education and health: the effects of quantity, credential, and selectivity. Demography 1999;36:445–60. [PubMed: 10604074]

20. White, S.; Dillow, S. Key Concepts and Features of the 2003 National Assessment of Adult Literacy (NAALS) (NCES 2006–471). Washington, DC: Department of Education, National Center for Education Statistics; 2005.

21. Botman SL, Moore TF, Moriarity CL, Parsons VL. Design and Estimation for the National Health Interview Survey (NHIS), 1995–2004, National Center for Health Statistics. Vital Health Stat 2 2000;130:1–31. [PubMed: 11707926]

22. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. J Clin Epidemiol 1988;41:1105–16. [PubMed: 3204420]

23. Ross CE, Wu CL. The links between education and health. Am Sociol Rev 1995;60:719–45.

24. Robert S, House JS. SES differentials in health by age and alternative indicators of SES. J Aging Health 1996;8:359–88. [PubMed: 10165980]

25. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000;23:943–50. [PubMed: 10895844]

26. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35:1381–95. [PubMed: 12900694]

27. Gordon AJ, Maisto SA, McNeil M, et al. Three questions can detect hazardous drinkers. J Fam Pract 2001;50:313–20. [PubMed: 11300982]

28. National Health and Nutrition Examination Survey (NHANES) Questionnaire (2001–2002). Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2002.

29. Behavioral Risk Factor Surveillance System (BRFSS). Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2003.

30. Chien-Wen T, Phillips RL, Green LA, Fryer GE, Dovey SM. What physicians need to know about seniors and limited prescription benefits, and why. Am Fam Physician 2002;66:212. [PubMed: 12152958]

31. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986;24:67–74. [PubMed: 3945130]

32. Barringer TA, Kirk JK, Santaniello AC, Foley KL, Michielutte R. Effect of a multivitamin and mineral supplement on infection and quality of life. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2003;138:365–71. [PubMed: 12614088]

33. Ross CE, Mirowsky J. Refining the association between education and health: the effects of quantity, credential, and selectivity. Demography 1999;36:445–60. [PubMed: 10604074]

34. Peyrot M, Rubin RR. Structure and correlates of diabetes-specific locus of control. Diabetes Care 1994;17:994–1001. [PubMed: 7988321]

35. Wagenknecht LE, Mayer EJ, Rewers M, et al. The insulin resistance atherosclerosis study (IRAS) objectives, design, and recruitment results. Ann Epidemiol 1995;5:464–72. [PubMed: 8680609]

36. King DE, Mainous AG III, Pearson WS. C-reactive protein, diabetes, and attendance at religious services. Diabetes Care 2002;25:1172–76. [PubMed: 12087015]

37. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 1979;109:186–204. [PubMed: 425958]

38. Ware, JE.; Kosinski, M.; Keller, SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Vol. 2. Boston: The Health Institute, New England Medical Center; 1995.

39. Ware, JE.; Kosinksi, M.; Dewey, JE.; Gandek, B. How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Lincoln, RI: QualityMetric Incorporated; 2001.

40. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24:385–96. [PubMed: 6668417]

41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–13. [PubMed: 11556941]

42. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. JAMA 1999;282:1737–44. [PubMed: 10568646]

43. The TRIAD Study Group. The translating research into action for diabetes (TRIAD) study: a multicenter study of diabetes in managed care. Diabetes Care 2002;25:386–89. [PubMed: 11815515]

44. Ayas NT, White DP, Al Delaimy WK, et al. A prospective study of self-reported sleep duration and incident diabetes in women. Diabetes Care 2003;26:380–84. [PubMed: 12547866]

45. Sickel AE, Moore PJ, Adler NE, Williams DR, Jackson JS. The differential effects of sleep quality and quantity on the relationship between SES and health. Ann NY Acad Sci 1999;896:431–34. [PubMed: 10681943]

46. Moore PJ, Adler NE, Williams DR, Jackson JS. Socioeconomic status and health: the role of sleep. Psychosom Med 2002;64:337–44. [PubMed: 11914451]

47. Benet-Martinez V, John OP. Los Cinco Grandes across cultures and ethnic groups: multitrait multimethod analyses of the Big Five in Spanish and English. J Pers Soc Psychol 1998;75:729–50. [PubMed: 9781409]

48. Srivastava S, John OP, Gosling SD, Potter J. Development of personality in early and middle adulthood: set like plaster or persistent change? J Pers Soc Psychol 2003;84:1041–53. [PubMed: 12757147]

49. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. J Res Pers 2003;37:504–28.

50. Schillinger D, Grumbach K, Piette J, et al. Association of health literacy with diabetes outcomes. JAMA 2002;288:475–82. [PubMed: 12132978]

51. Schillinger D, Barton LR, Karter AJ, Wang F, Adler N. Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes. Public Health Rep 2006;121:245–54. [PubMed: 16640146]

52. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004;36:588–94. [PubMed: 15343421]

53. The American Association of Clinical Endocrinologists. Medical Guidelines for the Management of Diabetes Mellitus: The AACE System of Intensive Diabetes Self-Management—2002 Update. Endocr Pract 2002;8(Suppl 1):40–82. [PubMed: 11939758]

54. Fitzgerald JT, Funnell MM, Hess GE, et al. The reliability and validity of a brief diabetes knowledge test. Diabetes Care 1998;21:706–10. [PubMed: 9589228]

55. Piette JD, Heisler M, Wagner TH. Problems paying out-of-pocket medication costs among older adults with diabetes. Diabetes Care 2004;27:384–91. [PubMed: 14747218]

56. Epidemiology of Diabetes Interventions and Complications (EDIC). Design, implementation, and preliminary results of a long-term follow-up of the diabetes control and complications trial cohort. Diabetes Care 1999;22:99–111. [PubMed: 10333910]

57. Krieger N, Sidney S, Coakley E. Racial discrimination and skin color in the CARDIA study: implications for public health research. Coronary artery risk development in young adults. Am J Public Health 1998;88:1308–13. [PubMed: 9736868]

58. Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. Soc Sci Med 2005;61:1576–96. [PubMed: 16005789]

59. Crofton, C. AHCPR Pub. No. 97-R079. Washington DC: Agency for Health Care Policy and Research; 1998. CAHPS 2.0 Questionnaires.

60. Hays RD, Shaul JA, Williams VS, et al. Psychometric properties of the CAHPS 1.0 survey measures. Consumer assessment of health plans study. Med Care 1999;37:MS22–31. [PubMed: 10098556]

61. Anderson LA, Dedrick RF. Development of the trust in physician scale: a measure to assess interpersonal trust in patient-physician relationships. Psychol Rep 1990;67:1091–100. [PubMed: 2084735]

62. Thom DH, Ribisl KM, Stewart AL, Luke DA. Further validation and reliability testing of the Trust in Physician Scale. The Stanford trust study physicians. Med Care 1999;37:510–17. [PubMed: 10335753]

63. Steiner JF, Koepsell TD, Fihn SD, Inui TS. A general method of compliance assessment using centralized pharmacy records. Description and validation. Med Care 1988;26:814–23. [PubMed: 3398608]

64. Clark DO, Von Korff M, Saunders K, Baluch WM, Simon GE. A chronic disease score with empirically derived weights. Med Care 1995;33:783–95. [PubMed: 7637401]

65. Kaiser Permanente Business Intelligence Report 2006. Research and Markets. 2006 [Accessed July 9, 2007]. Available at: http://www.researchandmarkets.com/reportinfo.asp?report_id=451423

66. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. Am J Public Health 1992;82:703–10. [PubMed: 1566949]

67. Hiatt RA, Friedman GD. Characteristics of patients referred for treatment of end-stage renal disease in a defined population. Am J Public Health 1982;72:829–33. [PubMed: 7091479]

68. Porter S, Jackson K, Trosclaire A, Pederson LL. Office on smoking and Health NCfCDPaHPC. Prevalence of current cigarette smoking among adults and changes in prevalence of current and some day smoking —United States, 1996—2001. MMWR 2003;52:303–7. [PubMed: 12731700]

69. Smith GD, Hart C, Blane D, Gillis C, Hawthorne V. Lifetime socioeconomic position and mortality: prospective observational study. BMJ 1997;314:547–52. [PubMed: 9055712]

70. Lynch JW, Kaplan GA, Shema SJ. Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. N Engl J Med 1997;337:1889–95. [PubMed: 9407157]

71. Melton LJ 3, Dyck PJ, Karnes JL, O'Brien PC, Service FJ. Non-response bias in studies of diabetic complications: the Rochester Diabetic Neuropathy Study. J Clin Epidemiol 1993;46:341–48. [PubMed: 8482998]

72. Robins JM. Association, causation, and marginal structural models. Synthese 1999;121:151–79.

73. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615–25. [PubMed: 15308962]

74. Bodnar LM, Davidian M, Siega-Riz AM, Tsiatis AA. Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology. Am J Epidemiol 2004;159:926–34. [PubMed: 15128604]

75. Bryan, JF.; Yu, Z.; van der Laan, M. U C Berkeley Division of Biolstatistics Working Paper Series. 2002. Analysis of longitudinal marginal structural models; p. 1-29.Paper 120

76. Neugebauer R, van der Laan M. Why prefer double robust estimators in causal inference? J Stat Plan Inference 2005;129:405–26.

77. Wang, Y.; Petersen, ML.; Bangsberg, D.; van der Laan, M. U C Berkeley Division of Biolstatistics Working Paper Series. 2006. Diagnosing bias in the inverse probability of treatment weighted estimator resulting from violation of experimental treatment assignment; p. 1-25.Paper 211

Table 1

Inclusion and exclusion criteria for the Kaiser Permanente Northern California Diabetes Registry[a]

Inclusion in the Registry is based on

1. Self-report from Kaiser member health surveys or

2. Pharmacy utilization of any insulin or oral hypoglycemic agents or

3. Laboratory results of glycosylated haemoglobin A1C (A1C) $\geq$ 7.0 or

4. Two or more abnormal glucose values (either fasting glucose $\geq$ 126 mg/dl or random glucose 5200 mg/dl) or

5. Outpatient utilization ($\geq$ 2 visits for diabetes) or

6. Inpatient hospital utilization (primary hospital discharge ICD9:250.xx).

Exclusion criteria are

1. Having no diabetes inclusion indicators during a sum total of two years of Kaiser membership after the date they were identified in the Registry or

2. Identified only by a ICD9 code of 648.8 (gestational diabetes) or

3. Identified due to pharmacy utilization of metformin or thiazolidinedione only (no other indicators) and also diagnosed with polycystic ovary syndrome, HIV lipodystrophy, metabolic syndrome, pre-diabetes or reproductive problems.

[a]The Registry is updated annually.

Table 2

Clinical, behavioral, psychosocial and quality and access indicators

Domain	Survey scales or variables
Demographic	Age
	Sex
	Race
	Nativity
	Language
Socioeconomic	Education
	• years of school
	• highest degree attained
	• quality-weighted attainment
	• functional health literacy
	Income
	Occupation
Clinical profile	Laboratory values
	Blood pressure
	Body mass index
	Inpatient and outpatient utilization
	Pharmacy utilization
	Type of diabetes
	Co-morbidity Index
	Self-reported clinical characteristics
Health behaviors, Adherence, Knowledge	Self-care behaviors
	• smoking
	• alcohol
	• exercise
	• medication refill adherence and discontinuation
	• self-monitoring of blood glucose
	• insulin injection frequency
	• appointment-keeping adherence
	• health screening adherence
	• sleep adequacy
	Diabetes knowledge
Psychosocial Characteristics	Big Five personality facets of conscientiousness and neuroticism
	Perceived stress
	Depressive and anxiety disorders
	Social support and social networks
	Discrimination
Socioeconomic Position (SEP)	Individual-level socioeconomic indicators

Domain	Survey scales or variables
	Neighbourhood SEP contextual variables
	Subjective measures of SEP
	Financial barriers
Patient-provider relationship	Patient-provider communication
	Primary care provider characteristics
Quality and access to care	Processes of care
	Referrals for specialty care

Table 3

Characteristics of survey respondents and non-respondents

	Respondents *n* (%) or Mean ± (SD)								Non-respondents
	African American	Asian (Chinese, Korean, Vietnamese, Japanese)	Caucasian	Filipino	Latino	Multiracial	Other[a] /unknown	All respondents	Non respondents Number (%) or Mean ± (SD)
n (row %)	3420 (16.9)	2312 (11.4)	4602 (22.8)	2404 (11.9)	3717 (18.4)	2222 (11.0)	1511 (7.5)	20 188 (100)	20 547 (100)
Survey Mode									
Telephone Interview	1806 (52.8)	971 (42.0)	2051 (44.6)	846 (35.2)	2229 (60.0)	1566 (70.5)	960 (63.5)	10429 (51.7)	0
Written-long	771 (22.5)	575 (24.9)	973 (21.1)	723 (30.1)	676 (18.2)	346 (15.6)	224 (14.8)	4288 (21.2)	0
Written-short	439 (12.8)	259 (11.2)	469 (10.2)	516 (21.5)	448 (12.0)	108 (4.9)	154 (10.2)	2393 (11.8)	0
Web	404 (11.8)	507 (21.9)	1109 (24.1)	319 (13.3)	364 (9.8)	202 (9.1)	173 (11.4)	3078 (15.2)	0
Female	1939 (56.7)	1019 (44.1)	2035 (44.2)	1246 (51.8)	1886 (50.7)	1122 (50.5)	593 (39.3)	9840 (48.7)	9537 (46.4)
Age (years)									
30–44	274 (8.0)	137 (5.9)	382 (8.3)	168 (7.0)	564 (15.2)	253 (11.4)	189 (12.5)	1967 (9.7)	2580 (12.6)
45–59	1347 (39.4)	934 (40.4)	1794 (39.0)	1106 (46.0)	1639 (44.1)	943 (42.4)	695 (46.0)	8458 (41.9)	8755 (42.6)
60+	1799 (52.6)	1241 (53.7)	2426 (52.7)	1130 (47.0)	1514 (40.7)	1026 (46.2)	627 (41.5)	9763 (48.4)	9212 (44.8)
Language preference English speaking	3221 (94.2)	1376 (59.5)	4394 (95.5)	1749 (72.8)	2217 (59.6)	1691 (76.1)	1164 (77.0)	15 812 (78.3)	17 212 (83.8)
Mean A1C ≥ 7	1655 (54.0)	1087 (50.1)	1915 (46.8)	1352 (60.4)	1957 (58.6)	1098 (56.2)	779 (57.9)	9843 (54.1)	8343 (51.3)
Mean LDL ≥ 100	1250 (41.7)	626 (28.9)	1391 (34.6)	634 (28.6)	1202 (37.1)	736 (38.1)	482 (36.3)	6321 (35.3)	6909 (42.4)
Uncontrolled blood pressure (S>130 or D>80)	1966 (61.4)	913 (42.8)	2264 (54.7)	1041 (46.5)	1746 (51.2)	1089 (54.2)	681 (49.5)	9700 (52.4)	9581 (54.3)
Missed appointment rate (per year)	0.17 ± (0.18)	0.10 ± (0.14)	0.13 ± (0.16)	0.13 ± (0.17)	0.18 ± (0.19)	0.17 ± (0.19)	0.18 ± (0.19)	0.15 ± (0.17)	0.18 ± (0.22)
Poor medication adherence[b]	955 (37.1)	461 (25.1)	961 (28.2)	592 (29.3)	1247 (42.8)	636 (36.4)	424 (35.1)	5276 (33.6)	5108 (39.0)
Smoking (current)	348 (10.2)	93 (4.0)	410 (9.0)	147 (6.1)	245 (6.6)	200 (9.1)	105 (7.0)	1548 (7.7)	1829 (9.0)
Practices SMBG	1573 (46.0)	1004 (43.4)	2246 (48.8)	1115 (46.4)	1622 (43.6)	1025 (46.1)	623 (41.2)	9208 (45.6)	6532 (31.8)
Outpatient risk score[c]	2.4 (2.1)	1.8 (1.1)	2.1 (1.7)	1.8 (1.0)	1.9 (1.4)	2.2 (1.6)	1.9 (1.4)	2.0 (1.6)	1.8 (1.5)
Inpatient risk score[c]	2.6 (4.2)	1.5 (2.0)	2.2 (3.3)	1.4 (2.1)	1.8 (3.0)	2.2 (3.9)	1.7 (2.9)	2.0 (3.2)	1.8 (3.4)
Linguistically isolated neighbourhood[d]	105 (3.1)	115 (5.0)	55 (1.2)	114 (4.8)	281 (7.6)	113 (5.1)	62 (4.1)	845 (4.2)	1035 (5.1)
Economically deprived neighbourhood[e]	827 (24.3)	124 (5.4)	346 (7.6)	148 (6.2)	668 (18.2)	323 (14.7)	167 (11.2)	2603 (13.0)	2724 (13.4)
Working class neighbourhood[f]	1440 (42.2)	401 (17.4)	1168 (25.8)	884 (37.0)	1749 (47.6)	868 (39.4)	464 (31.0)	6974 (34.9)	7566 (37.2)

[a] Includes Pacific Islanders, Native Americans, South Asian, other unspecified and unknown (did not answer).

[b] ≥20% continuous medication gaps.[63]

[c] They are weighted by patient's health care utilization and severity of disease.[64]

[d] ≥25% of households in census block group are linguistically isolated.

[e] ≥20% of households in census block group are living below poverty line.

[f] ≥66% of persons in census block group are employed in working class occupation.