PUBLISHED VERSION

Nasser Bagheri, Ian McRae, Paul Konings, Danielle Butler, Kirsty Douglas, Peter Del Fante, Robert Adams

http://hdl.handle.net/2440/90324

BMJ Open

Undiagnosed diabetes from cross-sectional GP practice data: an approach to identify communities with high likelihood of undiagnosed diabetes

Nasser Bagheri,[1] Ian McRae,[1] Paul Konings,[1] Danielle Butler,[1] Kirsty Douglas,[2] Peter Del Fante,[3] Robert Adams[4]

For numbered affiliations see end of article.

Correspondence to
Dr Nasser Bagheri;
nasser.bagheri@anu.edu.au

ABSTRACT

Objectives: To estimate undiagnosed diabetes prevalence from general practitioner (GP) practice data and identify areas with high levels of undiagnosed and diagnosed diabetes.

Design: Data from the North-West Adelaide Health Survey (NWAHS) were used to develop a model which predicts total diabetes at a small area. This model was then applied to cross-sectional data from general practices to predict the total level of expected diabetes. The difference between total expected and already diagnosed diabetes was defined as undiagnosed diabetes prevalence and was estimated for each small area. The patterns of diagnosed and undiagnosed diabetes were mapped to highlight the areas of high prevalence.

Setting: North-West Adelaide, Australia.

Participants: This study used two population samples—one from the de-identified GP practice data (n=9327 active patients, aged 18 years and over) and another from NWAHS (n=4056, aged 18 years and over).

Main outcome measures: Total diabetes prevalence, diagnosed and undiagnosed diabetes prevalence at GP practice and Statistical Area Level 1.

Results: Overall, it was estimated that there was one case of undiagnosed diabetes for every 3–4 diagnosed cases among the 9327 active patients analysed. The highest prevalence of diagnosed diabetes was seen in areas of lower socioeconomic status. However, the prevalence of undiagnosed diabetes was substantially higher in the least disadvantaged areas.

Conclusions: The method can be used to estimate population prevalence of diabetes from general practices wherever these data are available. This approach both flags the possibility that undiagnosed diabetes may be a problem of less disadvantaged social groups, and provides a tool to identify areas with high levels of unmet need for diabetes care which would enable policy makers to apply geographic targeting of effective interventions.

Strengths and limitations of this study

- The first study in Australia to examine undiagnosed diabetes at GP practice level with a reasonably large sample that allows us to describe patterns of undiagnosed diabetes in the population and also does enable highlighting of areas where active case finding may be warranted.
- The study illustrates a methodology which can be used as a tool to identify areas of high levels of unmet need for diabetes care. This could enable geographic targeting of effective interventions for enhancing early and timely detection and management of diabetes in those communities.
- The study shows that exploring the spatial variation in the pattern of diagnosed and undiagnosed diabetes can assist in identifying communities with high probability of having significant numbers of people with undiagnosed diabetes.
- Both the clinical data on which this study was based, and the survey on which the modelling was based, were drawn from a defined area within Adelaide. As such, while some of the conclusions can be generalised to the whole of Australia others are likely to be area specific and may not generalisable.

INTRODUCTION

Diabetes is a common chronic disease that significantly affects the health of people worldwide. It may lead to a range of complications which can cause disability and reduce the quality of life and life expectancy.[1 2] Diabetes constitutes a significant health and social burden in the community, and is one of the top 10 causes of death in Australia.[1] Lifestyle-related chronic diseases such as diabetes are predicted to rise rapidly over the next few decades worldwide,[3] posing challenges that will need to be met by effective preventive medicine strategies and health services planning.[1]

The study reported in this paper relates to data from an area in Western Adelaide, the capital city of South Australia, and much of the contextual data therefore relates to Australia. While the actual numbers identified in the study will be pertinent to Australia and particularly to Western Adelaide, the patterns identified and the methodology used will be relevant to studies of diabetes, and particularly of undiagnosed diabetes, undertaken anywhere in the world.

Currently, in Australia, approximately 875 400 (4.0%) people aged 15 years and over have been diagnosed with diabetes.[4] Some estimates suggest that this figure will rise to as much as 2 million by 2025 as a result of increasing obesity, ageing, and changes in ethnic composition of the Australian population.[5]

Already diagnosed prevalence and morbidity data underestimate the actual burden of diabetes because the disease is usually not diagnosed until it is clinically apparent. A number of local surveys[6] and national surveys[7] [8] have reported both diagnosed and undiagnosed diabetes based on population health surveys but the relative prevalence of diagnosed and undiagnosed diabetes reported varied widely among those studies. The North-West Adelaide Health Survey (NWAHS), for example, found a ratio of 5–6:1 for diagnosed versus undiagnosed diabetes,[9] which is consistent with the latest Australian Bureau of Statistics (ABS) National Health Survey data showing a ratio of 5:1,[10] whereas the earlier Australian Diabetes, Obesity and Lifestyle Study (AusDiab) estimated one undiagnosed case for every diagnosed cases of diabetes in Australia.[11]

The purpose of the current study was to revisit estimates of undiagnosed diabetes using a model based on NWAHS data applied to data from a large general practice, and to explore the nature of the areas with high levels of undiagnosed diabetes. This research broadly follows Nacul et al,[12] who used the Health Survey of England to estimate the small area prevalence of common obstructive pulmonary disease. They also compared expected model-based prevalence and observed prevalence in small areas (local authorities) in England using general practice data.

The main objective of this study was to explore patterns of undiagnosed diabetes in people aged 18 and over across small areas Statistical Area Level 1 (SA1s). The specific objectives were to (1) estimate the prevalence of total expected diabetes based on a model derived from the NWAHS data, and to compare this with the prevalence of general practice (GP)-diagnosed diabetes, (2) predict undiagnosed diabetes as the difference between total expected and already diagnosed diabetes, (3) identify associations between undiagnosed diabetes and socioeconomic status and (4) visualise the pattern of diagnosed and undiagnosed diabetes at SA1.

RESEARCH DESIGN AND METHODS

To estimate the expected level of diabetes in an area requires modelling the likelihood of diabetes based on all available relevant variables. Models using demographic data to estimate diabetes prevalence have been previously proposed,[11] [13–15] but they fail to take into account lifestyle factors and clinical measures.[16–18] Using data from the NWAHS[6] we developed a multivariate logistic regression model which includes demographic, lifestyle and clinical measures to estimate the probability of an individual having diabetes. This model was then applied to data from a large GP practice population in Western Adelaide to estimate the likely numbers of people with diabetes in various demographic and health categories, as well as in each SA1. Statistical areas are defined by the A B S[19] with an average population for SA1s of around 400 people. We compared these small area prevalence measures with the numbers in the GP practice population actually diagnosed with diabetes in each of these areas to estimate the numbers undiagnosed. These estimates were visualised on maps to explore the patterns of the areas with the highest levels of undiagnosed and diagnosed diabetes prevalence.

Data sources

This study used two datasets—one to develop a model of diabetes prevalence and another to apply this model to estimate the levels and location of diabetes. The first wave of the NWAHS is a representative population sample of people aged 18 years and over, living in the North-West region of Adelaide (n=4056). NWAHS established baseline self-reported and biomedically measured information which enabled them to identify all people with diabetes and not just the diagnosed. In the NWAHS, people with diagnosed diabetes were defined as those reporting that they had been told by a doctor that they had diabetes. This study was used to develop a model to estimate total expected diabetes from GP practice data.

The clinical data were extracted from a multisite GP practice in an area of Western Adelaide which includes the LeFevre Peninsula. People were defined as being diagnosed with diabetes if they had a diagnosis which was current at the time of the data extraction. While patients came from a wide area, they were concentrated in the Lefevre Peninsula area, and comprised 18% of the population of this well-defined area. Overall, 14 969 active patients were selected from the practice, of whom 12 271 were aged 18 years and over. As a substantial part of the analysis was based on geography, patients were classified according to the SA1 in which they resided. While the population of patients from the clinic was concentrated in the Lefevre Peninsula and nearby areas, there were patients from many other areas and to maintain confidentiality, the study also excluded those SA1s which comprised less than five patients (1564 patients) leading to a final sample of 10 707 people.

There was a range of missing values within the clinical dataset, the most important of which was body mass index (BMI) which was only available for 36% of the population. BMI was imputed where it was unavailable

using a regression imputation technique based on age, gender, receipt of pension, systolic and diastolic blood pressure and index of relative disadvantage.[17][20] The index of relative socioeconomic disadvantage (IRSD)[19] was a sub-index of the socioeconomic indexes for areas which were developed by the ABS.[19] The imputation model was estimated from the NWAHS data. As some of these regression variables were missing in the clinical data it was not possible to impute BMI for all patients; however, the imputation increased the BMI completeness rate from 36% to 88%. Thus, a final sample of 9327 patients was available with all the variables necessary for estimating total expected diabetes.

Model construction for total expected diabetes estimation

First, a set of risk and lifestyle factors including age, sex, smoking, cholesterol, BMI, systolic and diastolic blood pressure were identified based on previous studies[17][20] and availability in both the clinical dataset and the NWAHS dataset. Second, we included individual pension status and the area-based IRSD into the model to examine the relationship between socioeconomic status and diagnosed and undiagnosed diabetes. Owing to the high rate of missing values for cholesterol in the practice level data it was not included in the model. Finally, a multivariate logistic regression model including age, sex, smoking, pension status, systolic and diastolic blood pressure, BMI and IRSD was developed to predict the probability of an individual having diabetes in the GP practice population. The model was then applied to the GP practice dataset to estimate the total diabetes prevalence for the GP practice population in each relevant SA1. The difference between total expected and already diagnosed diabetes cases was calculated as undiagnosed diabetes prevalence both for the overall GP practice population and for age, gender and various other groups in practice population, which was also calculated at the SA1 level.

As we are comparing actual diagnoses with estimated overall prevalence, there is a risk of estimating negative levels of undiagnosed diabetes in categories where the diagnosis levels are high, for example among the obese where testing for diabetes would be expected to be almost universal. This does in fact occur, leading to some negative estimates.

In the final step, the area level prevalence was visualised for the SA1s within the Lefevre Peninsula to highlight areas with high and low prevalence of diagnosed and undiagnosed diabetes. The pattern of the IRSD was also mapped for the Lefevre Peninsula to allow visual comparison with the pattern of diagnosed and undiagnosed diabetes. Statistical analyses were carried out in Stata, V.12.1, and spatial analyses were undertaken in ArcGIS V.10.2. General practitioner practice data were extracted using Pen CAT clinical data extraction tool, and the GRAPHC G-Tag system[21] developed at the Australian Primary Health Care Research Institute at the Australian National University was used to geocode and geolink the GP practice data.

RESULTS

The logistic regression model showed that those people who are men, over 50 years old, with higher BMI and systolic blood pressure, more disadvantaged, pensioner and smoker have high odds of diabetes (see online supplementary table S1). The level of diagnosed diabetes in the NWAHS was 6.6% as reported by Grant et al,[6] with the level of undiagnosed diabetes 1%. The level of diagnosed diabetes reported for the 12 271 patients aged 18 and over in the clinical data extracted from a general practice for our study was 8.2%, which changed only marginally to 8.4% when SA1s with a small number of patients were removed. Only 1.5% of the group with incomplete data were diagnosed with diabetes, meaning that the final sample has a higher level of diagnosis than the clinical group over; this prevalence was 9.5% when the patients with incomplete data were removed from the GP practice dataset.

The prevalence of diagnosed, undiagnosed and total expected diabetes by age, gender, BMI, smoking, pension and IRSD in the GP practice population is presented in table 1. CIs for the predicted prevalences reflect the predictive accuracy of the logistic regression model.

The patterns of total expected, diagnosed, undiagnosed diabetes and IRSD in each SA1 in the Lefevre Peninsula were visualised in the study area. The IRSD scores are classified into tertiles for the practice population and labelled as most, moderate and least disadvantaged for the purposes of the modelling and prediction. The scores are classified as quintiles in the mapping to align with the classifications of the other maps.

Overall expected prevalence and diagnosed prevalence of diabetes

The overall estimated prevalence of diabetes in people aged 18 years and over at the GP practice level was 12.3% comprising 1147 of the 9327 active patients in the GP practice. Overall, the prevalence of total expected diabetes was higher in patients who were men, with higher BMI, most disadvantaged, ex-smokers, pensioners and those who were over 40 years old.

A total of 9.5% (or 886 of 9327) active patients have already been diagnosed with diabetes. The highest diagnosed rate was observed in people who were with higher BMI, over 45 years old, men, more disadvantaged, pensioner and ex-smoker.

Undiagnosed prevalence of diabetes

The prevalence of undiagnosed diabetes, measured by the difference between expected and diagnosed prevalence, was estimated to be 2.8% (261 of 9327) of active patients. This means that for every three and four patients diagnosed with diabetes in this general practice

Table 1 Distribution of diagnosed, undiagnosed and total expected prevalence of diabetes at GP practice level

	N (%)	Diagnosed prevalence, % (95% CI)	Undiagnosed prevalence, % (95% CI)	Estimated total prevalence, % (95% CI)
Age				
18–29	1243 (13.3)	1.0 (0.4 to 1.5)	3.5 (3.0 to 4.1)	4.5 (4.3 to 4.6)
30–39	1195 (12.8)	1.4 (0.7 to 2.0)	2.3 (1.6 to 3.0)	3.7 (3.5 to 3.9)
40–49	1452 (15.5)	3.3 (2.4 to 4.3)	3.6 (2.7 to 4.5)	7.0 (6.8 to 7.2)
50–59	2143 (23.0)	6.7 (5.7 to 7.8)	5.4 (4.4 to 6.5)	12.2 (11.9 to 12.5)
60–69	1568 (16.8)	16.6 (14.7 to 18.4)	2.8 (1.1 to 4.5)	19.4 (18.8 to 19.9)
≥70	1726 (18.5)	23.3 (21.3 to 25.3)	−0.9 (−2.8 to 0.9)	22.4 (21.8 to 22.8)
Gender				
Male	3783 (40.6)	11.7 (10.6 to 12.7)	4.1 (3.1 to 5.0)	15.8 (15.4 to 16.1)
Female	5544 (59.4)	8.0 (7.2 to 8.7)	2.1 (1.4 to 2.7)	10.0 (9.8 to 10.3)
BMI				
Underweight	78 (0.84)	5.1 (0.1 to 10.0)	2.1 (−2.7 to 6.9)	7.2 (5.9 to 8.6)
Normal	2042 (21.9)	4.5 (3.6 to 5.4)	4.1 (3.2 to 5.0)	8.6 (8.3 to 8.9)
Overweight	5302 (56.8)	6.9 (6.2 to 7.5)	3.9 (3.2 to 4.5)	10.8 (10.5 to 11.3)
Obese	1905 (20.4)	22.2 (20.4 to 24.1)	−1.2 (−2.9 to 0.5)	21.0 (20.4 to 21.6)
Smoking				
Current smoker	4581 (55.2)	8.7 (7.9 to 9.5)	3.1 (2.3 to 3.8)	11.8 (11.5 to 12.1)
Ex-smoker	2280 (27.4)	14.0 (12.6 to 15.5)	1.3 (−0.04 to 2.6)	15.3 (14.8 to 15.8)
Never-smoked	1444 (17.4)	7.6 (6.2 to 8.9)	3.5 (2.2 to 4.8)	11.1 (10.6 to 11.6)
Pension				
No	5457 (58.5)	4.2 (3.7 to 4.7)	3.9 (3.4 to 4.4)	8.1 (8.0 to 8.3)
Yes	3870 (41.5)	16.8 (15.7 to 18.0)	1.4 (0.3 to 2.5)	18.3 (17.9 to 18.6)
IRSD				
Most disadvantaged	3152 (33.8)	13.0 (11.8 to 14.2)	2.5 (1.4 to 3.6)	15.5 (15.1 to 15.9)
Moderately disadvantaged	3102 (33.2)	9.0 (8.0,10.0)	2.0 (1.1 to 3.0)	11.1 (10.8 to 11.4)
Least disadvantaged	3073 (33.0)	6.3 (5.4 to 7.1)	4.1 (3.2 to 4.9)	10.4 (10.1 to 10.7)
Total	9327 (100.0)	9.5 (8.8 to 9.9)	2.8 (2.5 to 3.1)	12.3 (11.6 to 12.9)

BMI, body mass index; IRSD, index of relative socioeconomic disadvantage.

community there would be one undiagnosed person with diabetes. The undiagnosed prevalence rate was higher in men, among people who live in the least disadvantaged areas and do not receive government pensions. The BMI of the undiagnosed is more likely to be in the normal and overweight category, and ex-smokers are less likely to have undiagnosed diabetes.

Spatial pattern of expected, diagnosed, undiagnosed diabetes and socioeconomic pattern

The decile of the IRSD[19] was mapped to compare with the pattern of undiagnosed and diagnosed diabetes at SA1 level. Figure 1 shows broadly the most disadvantaged population live in the Eastern part of the study area near the industrialised regions and the least disadvantaged live in the Western beach side areas.

The percentage of expected total diabetes is higher in the most disadvantaged areas, particularly, in the Eastern part of the study area (figure 2). The prevalence of expected total diabetes ranged from 4.6% to 23.1% across the SA1s in the study catchment.

The spatial pattern of the already diagnosed prevalence was broadly consistent with the pattern of expected total

level of diabetes in the study at SA1 level, in that the highest prevalence was seen in the Eastern part of the study area (figure 3). The prevalence of diagnosed diabetes varied from 0% to 37.5% in the selected SA1s in the study area.

The pattern of undiagnosed prevalence was different from the pattern of total expected and already diagnosed diabetes, with the highest levels of undiagnosed diabetes observed in the North-West and South-West parts of the Peninsula (figure 4). This rate ranged from 0% to 22.8% across SA1s in the study area.

As the number of patients in each SA1 varied, and in some cases, was quite small, both diagnosed and expected prevalence vary not only due to 'real' effects but also due to sample size effects. The overall rates of diagnosed diabetes ranged from 0% to 50% across the full set of SA1s, with rates of 50% only arising in SA1s with few patients of the participating general practice. The maps shown in figures 1–4 are restricted to the Lefevre Peninsula area of the general practice catchment where there is the highest penetration of this practice, and the numbers are larger and more stable.

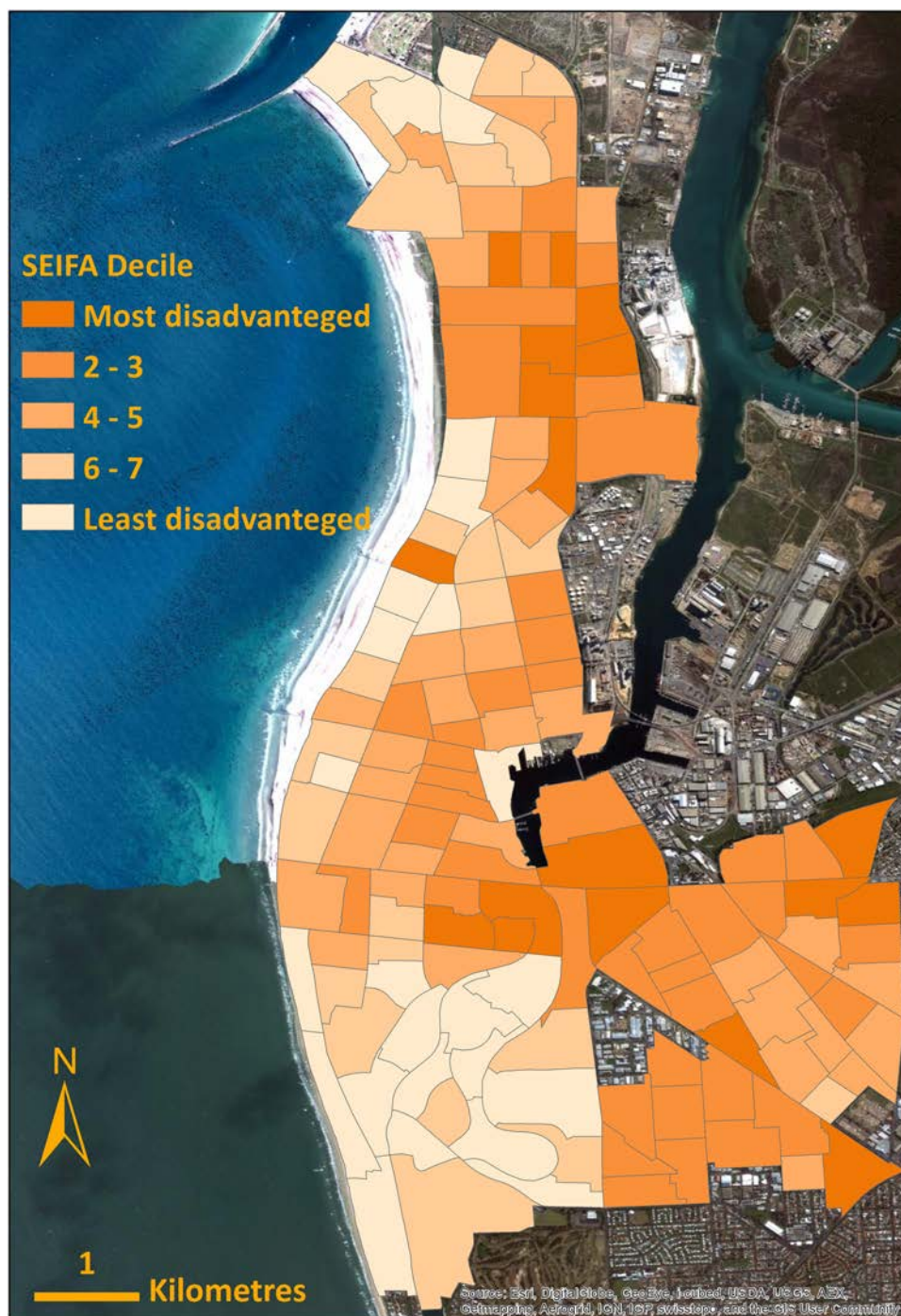Visual comparison of the spatial pattern of the expected, diagnosed and undiagnosed diabetes with the pattern of

Figure 1 Spatial patterns of the index of relative socioeconomic disadvantage.

IRSD shows that the prevalence of diagnosed diabetic patients is higher in the most disadvantaged areas and of undiagnosed cases higher in the least disadvantaged areas, consistent with the results shown in table 1.

DISCUSSION

We compared the measured diagnosed diabetes prevalence for a large multisite general practice in the Lefevre Peninsula and surrounding areas of Adelaide with the level predicted by a model derived from

NWAHS data. As would be expected, the prevalence of both diagnosed diabetes was closely related to BMI and to socioeconomic status (as measured by the IRSD). The level of undiagnosed diabetes, however, follows quite a different pattern, being rare among the obese and more likely among those living in the areas of least disadvantage. The latter effect is emphasised by the significant differences between those with and without pensions, where the pensioners are much more likely to have diagnosed diabetes and are much less likely to be undiagnosed. This will in part be due to age effects, as the
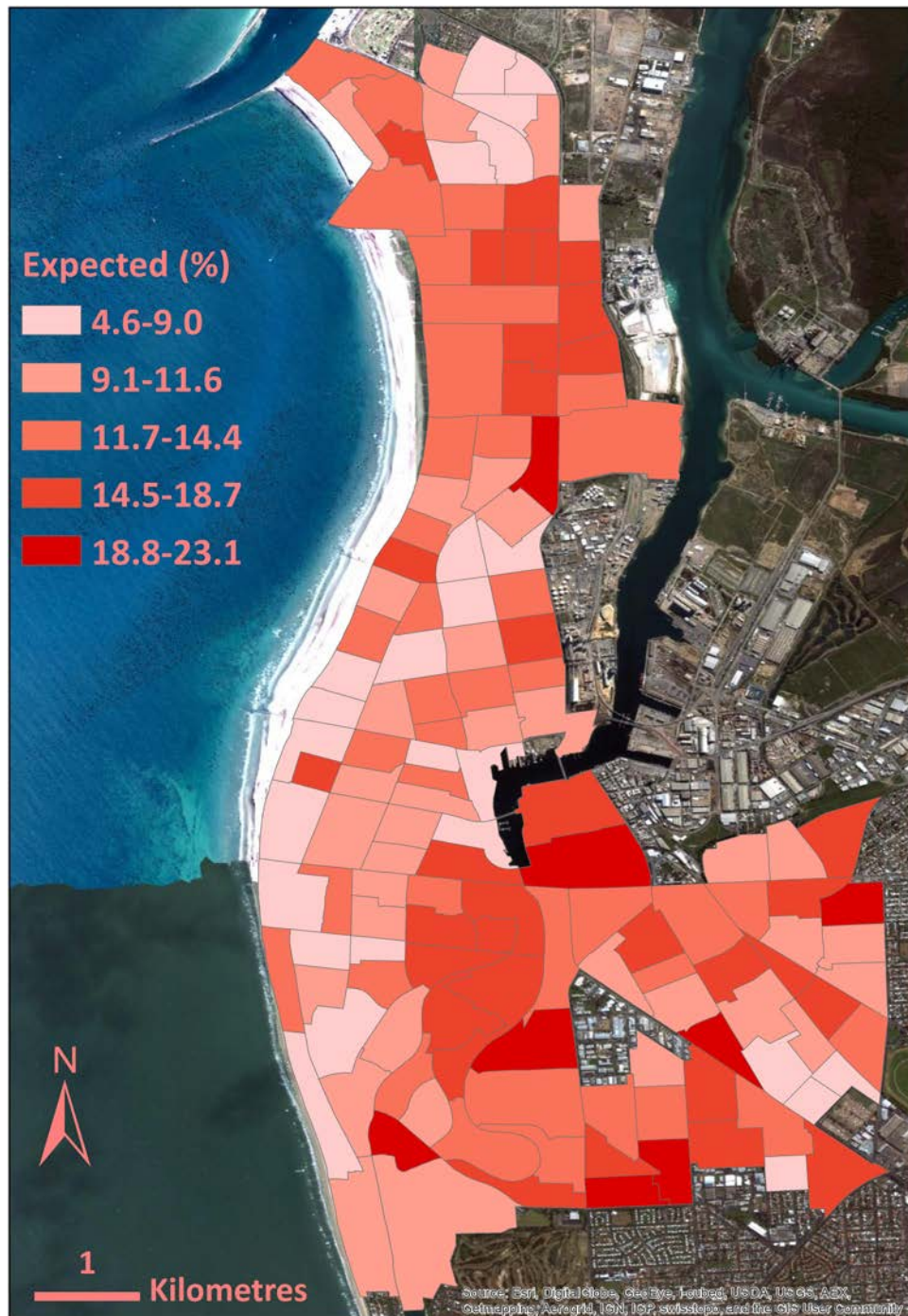
Figure 2 Spatial patterns of total expected diabetes.

diagnosed and total prevalence increase steadily with age, whereas for the undiagnosed group, while some differences are significant, there is no clear pattern except that the very old are unlikely to have undiagnosed diabetes.

The relationship between disadvantage and both diagnosed and undiagnosed diabetes is visible both in the overall estimates (table 1) and in the small area estimates reflected in the maps. The prevalence of total expected and diagnosed diabetes was high in the most socioeconomically disadvantaged areas, for example, in

the North-East part of the study area. However, the prevalence of undiagnosed diabetes was slightly higher in the least socioeconomically disadvantaged areas in the West part of the Peninsula.

The measures of diagnosed, undiagnosed and expected prevalence varied widely across the small SA1 areas, as would be expected with the relatively small samples in some of the areas; however, the patterns still stand out quite clearly.

This study found a ratio of 3–4:1 for diagnosed versus undiagnosed diabetes prevalence; in other words, for
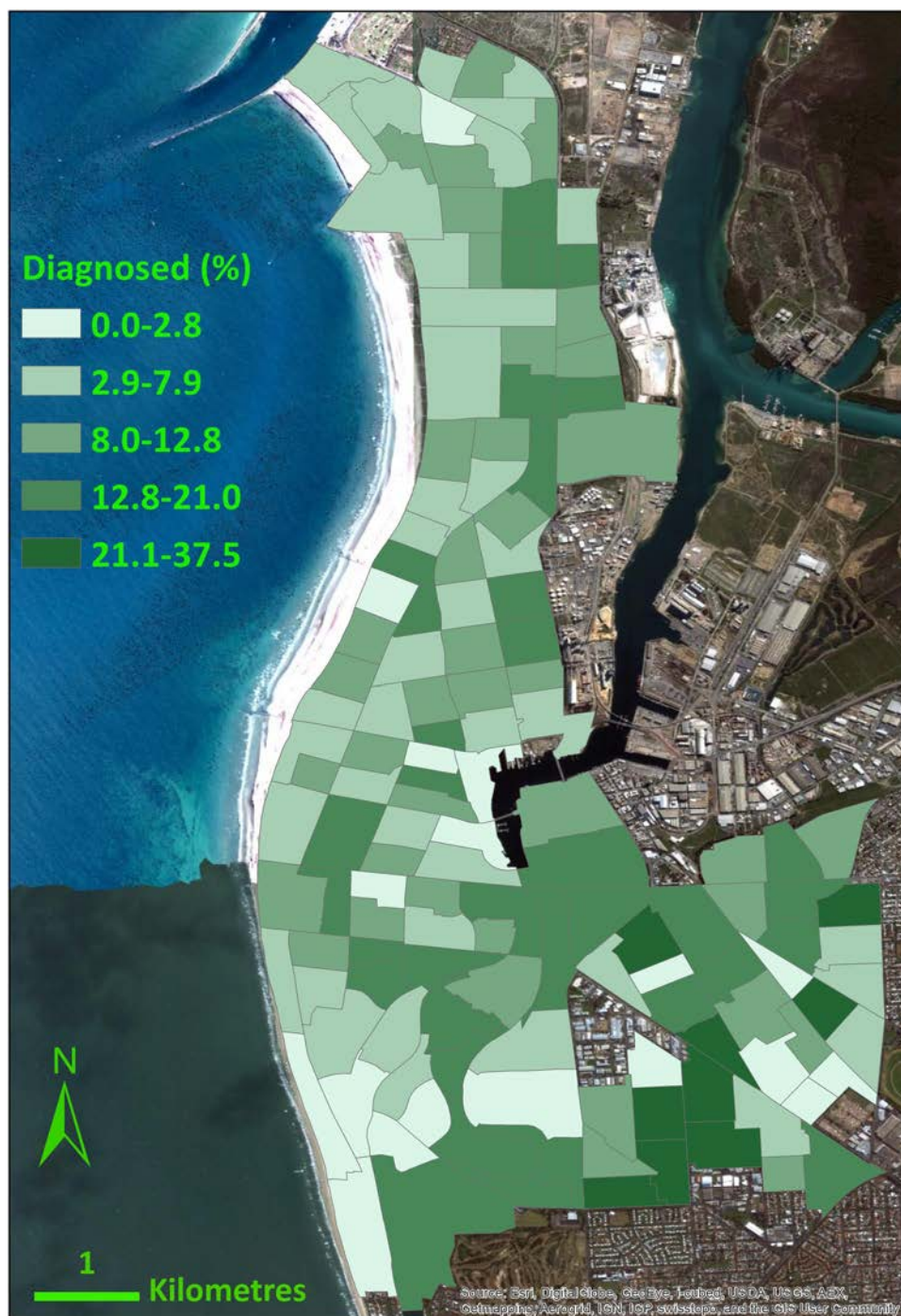
Figure 3 Spatial patterns of diagnosed diabetes.

approximately every three or four people with diagnosed diabetes, one person is likely to have undiagnosed diabetes. This ratio is of the same order as that of NWAHS which found a ratio of 1:5–6 for diagnosed versus undiagnosed,[6 9] although the differences may reflect the particular practice population which is likely to be of a somewhat lower socioeconomic status, and which has a relatively high level of diagnosed diabetes. Our findings, like the NWAHS findings and the most recent ABS study,[10] differ from the Australian Diabetes, Obesity and Lifestyle Study[22] which estimated that there was one

undiagnosed case for every diagnosed cases of diabetes in Australia.

The proportion of people with already diagnosed diabetes was estimated by the Australian Institute of Health and Wellbeing as 4.2% for Australia as a whole, and the NWAHS estimate for North-Western Adelaide was 6.6%. The higher levels of diagnosed diabetes found in the clinical practice used in this study (8.2% for the practice overall) may be due to the smaller area covered by this practice being of relatively low socioeconomic status, or to there being some selection of lower socioeconomic
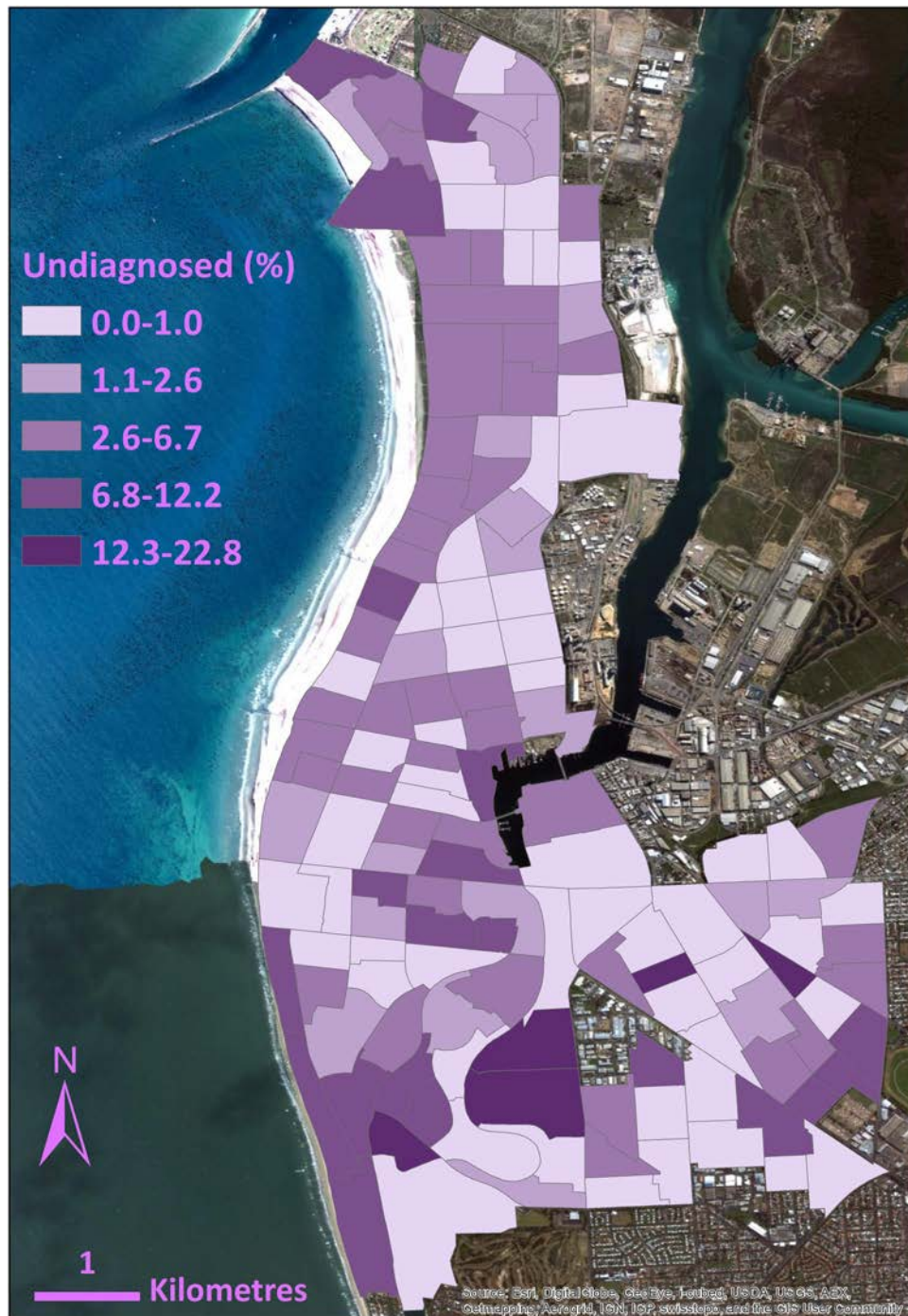
Figure 4 Spatial patterns of undiagnosed diabetes.

status patients to this practice. The higher 9.5% figure as noted above is due to the relatively low proportion of patients who were omitted due to having missing data having diagnosed diabetes. This would be expected, as most of those diagnosed would be regularly tested, and conversely, most of those who were not suspected of having diabetes would not be tested.

The patterns of our findings are consistent with Andersen et al[15] in the UK and Williams et al[23] studies in Australia, which found an association between area deprivation and diagnosed diabetes. Our findings which

showed a slightly higher rate of undiagnosed diabetes in areas with a low socioeconomic disadvantage are consistent with the Boston Area Community Health (BACH) study in the US A which confirms that undiagnosed diabetes is more prevalent in middle class, men and younger population.[24–28]

The BACH study also shows that GPs are more likely to diagnose diabetes from identical symptoms for those people who they believe are likely to have diabetes. In the USA, for example, undiagnosed signs and symptoms of type 2 diabetes in the community are patterned by

socioeconomic status rather than race/ethnicity, but following diagnosis by primary care physicians they are patterned more by race/ethnicity (rather than by socioeconomic status). They hypothesise that this apparent discrepancy is due to the fact that US physicians, probably unconsciously, attend to a patient's more visible racial and ethnic characteristics while overlooking less obvious socioeconomic status characteristics. Our findings suggest that a similar phenomenon may be occurring in the Lefevre Peninsula in Adelaide, not with race/ethnicity as the visible risk but with socioeconomic status and overweight/obesity. GPs may be overlooking less obvious diabetes risk factors, and/or not considering type 2 diabetes mellitus in those who do not fall within the obvious risk categories, most particularly of low socioeconomic status. When combined with the tendency of these people to not present or attend, as the higher socioeconomic status community is generally healthier, thus they remain undiagnosed.

Previous studies of undiagnosed diabetes have been based on surveys which collect both the patient's self-report information and blood tests to compare the self-report and the clinical state of the patients. Such studies are always constrained in size and geography. This study shows that investigation of undiagnosed diabetes at GP practice level with a reasonably large sample allows us to describe patterns of undiagnosed diabetes in the population and also does enable highlighting of areas where active case finding may be warranted. As such it provides a starting point for further studies which can use the breadth of data available in GP practices to vastly broaden the study of undiagnosed diabetes within Australia drawing on the previous surveys, but without the need for extremely large and expensive repeats of those surveys.

Moreover, this study shows that exploring the spatial variation in the pattern of diagnosed and undiagnosed diabetes, and estimating and visualising the prevalence of diagnosed and undiagnosed at area level, can assist in identifying communities with a high probability of having significant numbers of people with undiagnosed diabetes.

Both the clinical data on which this study was based, and the survey on which the modelling was based, were drawn from a defined area within Adelaide. As such, while some of the conclusions can be generalised to the whole of Australia, others are likely to be area specific and not generalisable. In particular, both the clinical data and NWAHS data find much higher levels of diagnosed diabetes than any of the available national studies.[10] This almost certainly reflects, in part, the socioeconomic status of the population within these areas; for example, 28% of the sample population live in SA1s in the lowest national quintile and 8% in the top quintile.

The clinical data on which the study was based included a higher proportion of people with diagnosed diabetes than the reported levels in the NWAHS. This may be due to further socioeconomic concentration for members of this particular general practice. The increase in diagnosed diabetes levels when patients with missing data are

removed from the sample suggests that few people with diagnosed diabetes are removed. The proportion of people with undiagnosed diabetes in the group who are deleted is of course not known, but as clinicians are more likely to test those people who they believe are at risk of diabetes it is probable that this group would have a relatively low risk of having undiagnosed diabetes.

The BMI, which is a well-known major determinant of diabetes risk, was not recorded for every individual in the GP practice level data and only 35.1% of active patients had a measured BMI in our sample data (3758 of 10 707). However, we imputed those BMI with missing values, based on a regression imputation model including all risk factors such as age, sex, smoking status, blood pressure, cholesterol level and pension in the model. While noting that there is a general rule of thumb that imputation should not be used for more than half a population, as regression imputation was used we were in effect using a synthetic estimate of the BMI where the original was not available, so while subject to variation the estimates were not subject to the biases potentially inherent in predictive mean matching or hot-decking methodologies.

CONCLUSIONS

The method can be used to estimate prevalence of diabetes at different geographical scales from a GP catchment to national level including demographics and clinical risk factors.

The approach taken in this study provides an opportunity for researchers who have access to general practice based clinical data to further explore prevalence, location and correlates of undiagnosed diabetes, and is applicable anywhere in the world where these data are available. This study flags the possibility that undiagnosed diabetes may be a problem of the less disadvantaged social groups, and illustrates a methodology which can be used as a tool to identify areas of high levels of unmet need for diabetes care. This could enable geographic targeting of effective interventions for enhancing early and timely detection and management of diabetes in those communities.

Author affiliations
[1]Australian Primary Health Care Research Institute, Australian National University, Canberra, Australia
[2]Department of General Practice, School of Medicine, Australian National University, Canberra, Australia
[3]Healthfirst Network, Adelaide, Australia
[4]School of Medicine, University of Adelaide, Adelaide, Australia

REFERENCES

1. AIHW. *Cardiovascular disease: Australian facts 2011*, in cardiovascular disease series no. 35. Cat. no. CVD 53. Canberra: AIHW, 2011.
2. Coppell KJ, Mann JI, Williams SM, *et al*. Prevalence of diagnosed and undiagnosed diabetes and prediabetes in New Zealand: findings from the 2008/09 Adult Nutrition Survey. *NZ Med J* 2013;126:25–42.
3. WHO. *The global burden of disease: 2004 update*. Geneva: World Health Organization, 2008.
4. ABS. Australian Health survey: First Results, 2011–12. 2012 [cited 3 October 2013]. http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/D4F2A67B76B06C12CA257AA30014BC65
5. Magliano DJ, Shaw JE, Shortreed SM. Lifetime risk and projected population prevalence of diabetes. *Diabetologia* 2008;51:2179–86.
6. Grant JF, Chittleborough CR, Taylor A, *et al*. The North West Adelaide Health Study: detailed methods and baseline segmentation of a cohort for selected chronic diseases. *Epidemiol Perspect Innov* 2006;3:4.
7. Barr E, Magliano D, Zimmet P, *et al*. *Tracking the accelerating epidemic: its causes and outcomes, in The Australian Diabetes, Obesity and Lifestyle Study*. Melbourne: International Diabetes Institute, 2005.
8. Diabetes Australia. What is diabetes? 2011 [cited 20 August 2013]. http://www.diabetesaustralia.com.au/en/Understanding-Diabetes/What-is-Diabetes/
9. Grant JF, Taylor AW, Ruffin RE, *et al*. Cohort profile: the North West Adelaide Health Study (NWAHS). *Int J Epidemiol* 2009;38:1479–86.
10. ABS. Australian Health Survey: Biomedical Results for Chronic Diseases, 2011–12. 2013 [cited 2013]. http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.005Chapter1002011-12
11. Dunstan DW, Zimmet PZ, Welborn TA, *et al*. The rising prevalence of diabetes and impaired glucose tolerance: the Australian diabetes. *Obesity and Lifestyle Study Diabetes Care* 2002;25:829–34.
12. Nacul L, Soljak M, Samarasundera E, *et al*. COPD in England: a comparison of expected, model-based prevalence and observed prevalence from general practice data. *J Public Health* 2011;33:108–16.
13. Appuhamy JA, Kebreab E, France J. A mathematical model for determining age-specific diabetes incidence and prevalence using body mass index. *Ann Epidemiol* 2013;23:248–54.
14. Boutayeb A, Chetouani A. A critical review of mathematical models and data used in diabetology. *Biomed Eng Online* 2006;5:1–9.
15. Andersen AF, Carson C, Watt HC, *et al*. Life-course socio-economic position, area deprivation and type 2 diabetes: findings from the British Women's Heart and Health Study. *Diabet Med* 2008;25:1462–8.
16. Tabaei B, Herman W. A multivariate logistic regression equation to screen for diabetes development and validation. *Diabetes Care* 2002;25:1999–2003.
17. Baan CA, Ruige JB, Stolk RP, *et al*. Performance of a predictive model to identify undiagnosed diabetes in a health care setting. *Diabetes Care* 1999;22:213–9.
18. Jaffiol C, Thomas F, Bean K, *et al*. Impact of socioeconomic status on diabetes and cardiovascular risk factors: results of a large French survey. *Diabetes Metab* 2013;39:56–62.
19. ABS. Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011 (cat. no. 2033.0.55.001). 2011 [cited 20 September 2013]. http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifa
20. Pierce MB, Zaninotto P, Steel N, *et al*. Undiagnosed diabetes—data from the English longitudinal study of ageing. *Diabet Med* 2009;26:679–85.
21. GRAPHC. G-Tag Systemm overview. 2013 [cited 2013]. http://graphc.aphcri.anu.edu.au/documents/G-Tag%20System%20Overview_v0.2.pdf
22. Diabetes Institute. Diabesity & Associated Disorders in Australia—2000. The Accelerating Epidemic. *The Australian Diabetes, Obesity and Lifestyle Study* (AusDiab). 2001 [cited 21 August 2013]. http://www.health.gov.au/internet/main/publishing.nsf/Content/pq-diabetes-pubs-diabesity
23. Williams ED, Magliano DJ, Zimmet PZ, *et al*. Area-level socioeconomic status and incidence of abnormal glucose metabolism the Australian Diabetes, Obesity and Lifestyle (AusDiab) study. *Diabetes Care* 2012;35:1455–61.
24. Piccolo R, McKinlay J, Marceau L. OP66A profile of undiagnosed diabetics in the community: results from the Boston Area Community Health (Bach) pre-diabetes survey (Abstract). *J Epidemiol Community Health* 2012;66:A26.
25. Rathmann W, Haastert B, Icks A, *et al*. Sex differences in the associations of socioeconomic status with undiagnosed diabetes mellitus and impaired glucose tolerance in the elderly population: the KORA Survey 2000. *Eur J Public Health* 2005;15:627–33.
26. Zhang X, Beckles GL, Bullard KM, *et al*. Access to health care and undiagnosed diabetes along the United States-Mexico border. *Rev Panam Salud Publica* 2010;28:182–9.
27. McKinlay J, Chiu G. *Does a serendipitous diagnosis differentiate diagnosed from undiagnosed diabetes? Results from a longitudinal investigation. In 2013 Academy Health Annual Research Meeting*. Baltimore: New England Research Institute, 2013.
28. McKinlay J, Marceau LD, Piccolo RJ. Do doctors contribute to the social patterning of disease: the case of race/ethnic disparities in diabetes mellitus. *Med Care Res Rev* 2012;69:176–93.