Canadian Journal on Aging / La Revue canadienne du vieillissement

http://journals.cambridge.org/CJG

Validating Chronic Disease Ascertainment Algorithms for Use in the Canadian Longitudinal Study on Aging*

Mark Oremus, Ronald Postuma, Lauren Griffith, Cynthia Balion, Christina Wolfson, Susan Kirkland, Christopher Patterson, Harry S. Shannon and Parminder Raina

Validating Chronic Disease Ascertainment Algorithms for Use in the Canadian Longitudinal Study on Aging*

Mark Oremus,[1] Ronald Postuma,[2] Lauren Griffith,[1] Cynthia Balion,[3] Christina Wolfson,[4,5] Susan Kirkland,[6,7] Christopher Patterson,[8] Harry S. Shannon,[1] CLSA Algorithms Validation Group,♦ and Parminder Raina[1]

RÉSUMÉ

Nous avons validé sept algorithmes d'évaluation de maladie chronique pour l'usage dans L'Étude longitudinale canadienne (ÉLCV) sur le vieillissement. Les algorithmes ont concerné le diabète type 2, parkinsonisme, obstruction chronique de flux d'air, ostéoarthrite de main, ostéoarthrite de hanche, ostéoarthrite de genou, et la maladie cardiaque ischémique. Notre recrutement de cible était 20 cas et contrôles par chaque maladie. Quelques cas ont été utilisés comme contrôles avec certaines maladies. Tous les participants ont répondu à des questionnaires au sujet des symptômes de la maladie et d'utilisation de médicaments. Les cas et les contrôles de diabète ont subi le test de jeûne de glucose et les cas et les contrôles de l'obstruction chronique de flux d'air ont subi le test de spirométrie. Pour chaque maladie, nous avons utilisé l'algorithme adapté pour classifier si les participants étaient positifs ou négatif pour la maladie. Nous avons également calculé la sensibilité et la spécificité utilisant le diagnostic de médecin comme norme. L'échantillon final a fait participer 176 participants, qui ont été recrutés dans trois villes canadiennes entre 2009 et 2011. La plupart des sensibilités et spécificités étaient 80% ou plus, indiquant que les sept algorithmes peuvent correctement identifier des personnes avec les maladies.

ABSTRACT

We validated seven chronic disease ascertainment algorithms for use in the Canadian Longitudinal Study on Aging. The algorithms pertained to diabetes mellitus type 2, parkinsonism, chronic airflow obstruction (CAO), hand osteoarthritis (OA), hip OA, knee OA, and ischemic heart disease. Our target recruitment was 20 cases and controls per disease; some cases were controls for unrelated diseases. Participants completed interviewer-administered disease symptom and medication use questionnaires. Diabetes cases and controls underwent fasting glucose testing; CAO cases and controls underwent spirometry testing. For each disease, the appropriate algorithm was used to classify participants' disease status (positive or negative for disease). We also calculated sensitivity and specificity using physician diagnosis as the reference standard. The final sample involved 176 participants recruited in three Canadian cities between 2009 and 2011. Most estimated sensitivities and specificities were 80 per cent or more, indicating that the seven algorithms correctly identified individuals with the target disease.

[1] Department of Clinical Epidemiology and Biostatistics, McMaster University

[2] Department of Neurology, McGill University and Montréal General Hospital

[3] Department of Pathology and Molecular Medicine, McMaster University

[4] Department of Epidemiology and Biostatistics and Occupational Health, McGill University

[5] Research Institute, McGill University Health Centre

[6] Department of Community Health and Epidemiology, Dalhousie University

[7] Department of Medicine, Dalhousie University

[8] Geriatric Services, Hamilton Health Sciences and Division of Geriatric Medicine, McMaster University

♦ Jonathan D. Adachi, M.D. (McMaster University), Wade Elliot, M.D. (Regional Joint Assessment Centre, Hamilton, Ontario), Paul Hernandez, M.D. (Dalhousie University), Ally P.H. Prebtani, M.D. (McMaster University), Malcolm Sears, M.D. (McMaster University), Rolf J. Sebaldt, M.D. (McMaster University), Koon Teo, M.D. (McMaster University), Dale Williams, M.D. (McMaster University), James Williams, M.D. (Dundas Family Medical Group, Dundas, Ontario), Mitchell Winemaker, M.D. (McMaster University).

Correspondence and requests for offprints should be sent to / La correspondance et les demandes de tirés-à-part doivent être adressées à:

Parminder Raina, Ph.D.
McMaster University
1280 Main Street West
MIP-309A
Hamilton, ON L8S 4K1
(praina@mcmaster.ca)

The Canadian Longitudinal Study on Aging (CLSA) is a 20-year prospective study recruiting 50,000 persons between the ages of 45 and 85 years at baseline. All participants provide self-reported data on socio-demographic characteristics, lifestyles and behaviors, physical and clinical measures, psychological measures, economic measures, health status measures, and health services utilization. Of the 50,000 participants, 30,000 (i.e., the "CLSA Comprehensive") provide additional information via physical examination and biospecimen collection (Raina et al., 2009a). The CLSA tracks the natural history of chronic diseases and investigate the associations between a multitude of risk factors and the incidence of these diseases (Raina et al., 2009b).

The CLSA asks all 50,000 participants to self-report whether a health professional diagnosed them with any of 38 chronic conditions, including respiratory, cardiovascular, neurological, gastrointestinal, rheumatic, mental health, cancer, and vision-related conditions. For the 30,000 participants in the CLSA Comprehensive, self-reported diagnoses are supplemented with disease-specific questionnaires, physical test measures, and medication use data.

Evidence suggests that self-reported diagnoses have low accuracy for identifying many chronic diseases (Kriegsman, Penninx, van Eijik, Boeke, & Deeg, 1996), and participant assessment by health professionals is not feasible in the CLSA because of standardization difficulties and logistical constraints. Therefore, the CLSA decided to employ disease ascertainment algorithms to identify the presence of chronic diseases in the

Comprehensive participants. The algorithms combine outcomes from the self-reported diagnoses questions, disease-specific questionnaires, physical test measures, and medication usage data to classify participants into one of three general categories: diseased, possibly or probably diseased (sometimes referred to as uncertain), or not diseased.

The CLSA established a Clinical Working Group to identify published algorithms for use in the study (Raina et al., 2009a; Raina et al., 2009b). Through a systematic review of the literature (Raina et al., 2009b), the group found algorithms with evidence of concurrent validity for all except seven chronic conditions – diabetes mellitus type 2 (diabetes); parkinsonism; chronic airflow obstruction (CAO); osteoarthritis (OA) of the hand, OA of the hip, and OA of the knee; and ischemic heart disease (IHD). The Clinical Working Group developed algorithms for these seven conditions through a process of consensus discussion among group members. Existing diagnostic or disease management algorithms, clinical guidelines, and CLSA data collection instruments served as the basis for discussion. We conducted this pilot study to assess the validity of the seven algorithms against physician diagnosis.

Methods
Algorithms

The algorithms are shown in the supplemental online content as eFigures 1a–g (available at www.journals.cambridge.org/cjg2013002) and are briefly described here.

Diabetes

All participants who report taking medications for diabetes are classified as having the disease in this algorithm (see eFigure 1a). Participants who do not report taking diabetes medications but for whom fasting blood glucose levels are ≥ 7.0 mmol/l are classified as having diabetes. Participants with levels 6.1 to 6.9 mmol/l are classified as having impaired fasting glucose, while participants with levels < 6.1 mmol/l are classified as not having diabetes (Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2008).

Glucose was measured with various methods depending on where the patients had their blood collected (see the table in the supplemental online content file eAppendix A). The diagnostic cut points apply to all glucose methods performed in the laboratory.

Parkinsonism

The parkinsonism algorithm is based on questionnaire and medication information (see eFigure 1b). A self-reported diagnosis of Parkinson's disease, combined with a report of taking Parkinson's disease medications, leads to a classification of "probable parkinsonism". Self-reported diagnosis, without report of taking medications, or no self-reported diagnosis at all, will lead to assessment with the nine-item Tanner Questionnaire (Duarte et al., 1995).

The Tanner Questionnaire asks about the presence of disease symptoms such as shaking, poor balance, or "freezing" (becoming motionless) in doorways. Response options are dichotomous (yes/no); "yes" responses are assigned a value of 1, and "no" responses are assigned a value of 0. Scores < 3 indicate no parkinsonism, a score of 3 indicates possible or unconfirmed parkinsonism, and scores ≥ 4 indicate probable parkinsonism.

Although the self-report and medication questions, and the Tanner Questionnaire, ask about Parkinson's disease, a diagnosis of Parkinson's disease cannot be made without a clinical examination. Consequently, we deemed the algorithm most appropriate to ascertain parkinsonism, for which Parkinson's disease is the most common cause.

Chronic Airflow Obstruction (CAO)

The CAO algorithm includes self-report questions on the presence of symptoms for chronic obstructive pulmonary disease (COPD) or asthma (see eFigure 1c). In the absence of a complete clinical assessment, the CLSA cannot clearly distinguish between COPD and asthma, so the two conditions are combined into an entity called CAO. The CAO algorithm includes consideration of the FEV_1/FVC (forced expiratory volume in one second/forced vital capacity) ratio, which is derived from the spirometry pulmonary function test. Normal and abnormal FEV_1/FVC cut-off ratios for each participant are determined in accordance with age- and sex-specific reference values developed from a sample of 7,429 asymptomatic and non-smoking persons from the United States (Hankinson, Odencrantz, & Fedan, 1999).

Participants who self-report "no" to COPD or asthma symptoms and have normal-range FEV_1/FVC ratios are considered non-diseased, regardless of medication use. Participants who report symptoms and have normal range FEV_1/FVC ratios are classified as "possible CAO", regardless of medication use. An abnormal FEV_1/FVC ratio, irrespective of symptoms but with no report of medication use, also results in a classification of "possible CAO"; an abnormal ratio with a positive report of medication use is classified as "definite CAO". In the case of participants who self-report "yes" to the symptoms of COPD or asthma, more algorithm pathways lead to "definite CAO" to reflect the importance of a positive self-report.

Osteoarthritis (OA)

As noted, we developed separate algorithms for hand, hip, and knee OA (see eFigures 1d, 1e, and 1f). These algorithms involve queries about self-reported diagnoses, joint enlargement, hand or joint pain, hand enlargement, groin or thigh pain, hip or knee replacement, or pain or swelling in the knees. Various combinations of answers to these questions determine whether the presence of disease is definite, probable, possible, asymptomatic, or uncertain.

Ischemic Heart Disease (IHD)

The IHD algorithm combines myocardial infarction and angina pectoris into a single disease entity (see eFigure 1g). The algorithm contains a series of questions about self-reported diagnosis and symptoms, prior medical procedures, and medication use. The algorithm also uses the results of an electrocardiogram (ECG) and the Rose Questionnaire (Rose, McCartney, & Reid, 1977).

The Rose Questionnaire contains nine questions about the presence of pain or discomfort in the chest, the location of the pain, and whether pain persists when walking, standing still, or going uphill. Participants reporting pain or discomfort in the chest are considered as "positive" on the Rose Questionnaire if the pain is in the sternum, left arm, or left anterior chest; if they stop or slow down in response to pain while walking; if they indicate the pain gets better while standing still; and if they report pain duration of less than 10 minutes. The components of the algorithm are combined to ascertain whether participants have definite, probable, uncertain, or no disease.

Subject Recruitment and Study Process

Recruitment of validation study participants took place between June 2009 and June 2011. All participants were between ages 45 and 85 and spoke English or French.

We recruited 20 cases for six of the seven disease entities. Due to a lengthy recruitment time, we halted OA hip recruitment at 16 cases. Our initial intent was to validate only the diabetes, parkinsonism, and CAO algorithms. We planned to have diabetes cases serve as CAO controls, parkinsonism cases serve as diabetes controls, and CAO cases serve as parkinsonism controls. We subsequently added OA and IHD to the study and recruited 20 persons without OA to serve as controls for the OA algorithm and 20 persons without IHD to act as controls for the IHD algorithm. The 20 OA controls were used as comparators for all three OA algorithms.

We enrolled cases from specialized medical clinics in three Canadian cities (Hamilton, Montréal, and Halifax). Cases were eligible if they had physician-diagnosed disease for which they were receiving treatment. We excluded diabetes cases that had CAO, parkinsonism cases that had diabetes, and CAO cases that had parkinsonism. Physician-collaborators in the clinics consulted patient charts to identify cases that were free of the disease for which they would serve as a control.

We used advertisements to recruit OA controls from among McMaster University employees and patients in a Hamilton family practice clinic. We recruited IHD controls from non-diseased persons who were undergoing ECG exercise stress tests at two Hamilton medical clinics.

In the clinics, research assistants approached potential participants in waiting rooms and explained the study. They administered informed consent to persons who verbally agreed to participate. Following consent, the research assistants conducted face-to-face interviews that featured disease symptom and medication use questions (available from authors upon request). Diabetes cases and controls underwent fasting blood glucose testing, and CAO cases and controls underwent spirometry testing. IHD cases and controls underwent ECG unless their charts contained the results of an ECG ordered within the past 12 months.

The Hamilton Health Sciences/Faculty of Health Sciences Research Ethics Board approved the study. We also obtained ethics approval from the research ethics boards of St. Joseph's Healthcare, Hamilton; McGill University Health Centre, Montréal; and Capital Health, Halifax. All participants gave written informed consent prior to enrolment in the study.

Statistical Analysis

We estimated sensitivity and specificity for each algorithm by classifying all participants as "test positive" or "test negative" and using participants' case or control status as the reference standard. We utilized the Wilson score interval to compute a 95 per cent binomial proportion confidence interval for each sensitivity or specificity estimate. Estimates were rounded to the nearest integer. We used SAS v9.2 (SAS Institute, Cary, NC, United States) and OpenEpi v2.3.1 (www.OpenEpi.com) to conduct the analysis.

Although the algorithms identified some participants as diseased (classified as test positive) and other participants as non-diseased (classified as test negative), classifying all participants as test positive or test negative was a challenge. Data combinations in the algorithms did not permit all participants to be identified as diseased or non-diseased, so some participants were identified as possibly or probably diseased, or their disease status was labeled as "uncertain". We classified participants in the possibly or probably diseased, and uncertain, categories as test positive and calculated sensitivities and specificities. We then reclassified these participants as test negative and redid the sensitivity and specificity calculations. We conducted this reclassification for each algorithm.

Results

We recruited 176 participants who had a median age of 66 years (25th percentile = 55 years; 75th percentile = 73 years); 55 per cent (n = 97) were female, and 77 per cent (n = 136) reported being in good or very good overall health. Table 1 describes participant characteristics.

Tables 2a–e show the estimated sensitivities and specificities for each disease ascertainment algorithm. Out of 30 estimated sensitivities, 26 of the estimated sensitivities were 80 per cent or more; out of 30 estimated specificities, 22 of the estimated specificities were 80 per cent or more.

The diabetes algorithm demonstrated good overall performance. The optimal results occurred when we reclassified the following two groups of persons as test negative: (1) persons who were initially classified on the algorithm as having probable diabetes; (2) persons who were initially classified as having impaired fasting glucose. Specificity was lower than sensitivity, thereby suggesting the algorithm would produce more false positives than false negatives.

The parkinsonism algorithm's results were excellent. When we reclassified possible or unconfirmed parkinsonism on the algorithm as test negative, sensitivity and specificity were both 100 per cent.

Table 1: Participant characteristics (*n* = 176)

Characteristic	
Age (years)	
Median	66
25th – 75th percentile	55 – 73
Minimum	45
Maximum	85
Sex (%)	
Female	55
Male	45
Education (%)	
Elementary school	7
High school	32
College	24
University	36
Missing	1
Household Income (%)	
$0 – $30,000	16
$30,001 – $60,000	24
$60,001 or more	49
Missing	11
Overall health (%)	
Excellent	12
Very good	35
Good	42
Fair	8
Poor	3

For the CAO algorithm, large fluctuations in specificity were evident depending on how we reclassified possible CAO. The optimal reclassification was to consider persons with possible CAO as test negative and include the medication question in the algorithm (100% sensitivity, 80% specificity). Specificity rose to 90 per cent when we removed the medication question, but sensitivity decreased to 65 per cent. The CAO algorithm may detect most or all cases, but several non-diseased persons may test positive.

The OA hand algorithm performed best when uncertain or possible OA were reclassified as test positive. Under this reclassification scheme, the algorithm would detect most or all cases and have a very low false positive rate (i.e., 1 in 20 controls would test positive for hand OA).

Results for the OA hip algorithm involved several reclassifications of uncertain or probable OA, as well as testing the redundancy of a question about limitations in range of hip motion. Optimal algorithm performance (100% sensitivity, 95% specificity) occurred following reclassification of uncertain outcomes as test positive and probable OA as test negative. These results were unchanged after removing the hip motion question from the algorithm.

Turning to OA knee, the algorithm performed best when uncertain or probable OA were both reclassified as test positive (100% sensitivity, 95% specificity).

For the IHD algorithm, sensitivity remained constant and specificity decreased following the inclusion of ECG results. The inclusion of ECG led to misclassification of two controls with positive Q waves as test positive. The algorithm did not account for the fact that positive Q waves could indicate other conditions besides IHD. For example, patients with left ventricular hypertrophy may have S waves that look like Q waves. The optimal result on the IHD algorithm was 100 per cent sensitivity and 85 per cent specificity following reclassification of probable IHD as test positive and uncertain outcomes as test negative in the algorithm without ECG. Reclassifying probable IHD as test negative in the same algorithm, while keeping uncertain outcomes as test negative, increased specificity to 95 per cent yet decreased sensitivity to 80 per cent.

Discussion

We validated seven chronic disease algorithms for use in the CLSA. This validation was necessary because the seven algorithms were developed by the CLSA's Clinical Working Group. These algorithms had not been previously used in research studies. The other algorithms employed in the CLSA already had existing evidence of validity and were not examined in this pilot study.

The seven algorithms generally performed better at detecting persons with disease than identifying persons without disease. Each algorithm had at least one combination of reclassifications where sensitivity was 100 per cent. However, only the parkinsonism algorithm had a combination where sensitivity and specificity were both 100 per cent.

Table 2a: Diabetes algorithm sensitivities and specificities

Reclassification	Sensitivity (%) (95% Confidence Interval)	Specificity (%) (95% Confidence Interval)
IFG = test negative and probable diabetes = test negative	95 (76–99)	90 (70–97)
IFG = test positive and probable diabetes = test positive	100 (84–100)	70 (48–85)
IFG = test positive and probable diabetes = test negative	95 (76–99)	75 (53–89)
IFG = test negative and probable diabetes = test positive	100 (84–100)	85 (64–95)

IFG = impaired fasting glucose

Table 2b: Parkinsonism algorithm sensitivities and specificities

Reclassification[a]	Sensitivity (%) (95% Confidence Interval)	Specificity (%) (95% Confidence Interval)
Possible or unconfirmed parkinsonism = test positive	100 (84–100)	95 (76–99)
Possible or unconfirmed parkinsonism = test negative	100 (84–100)	100 (84–100)

[a] **Probable parkinsonism = test positive**

The main advantage of the algorithms is the ability to identify diseased and non-diseased study participants without a physician examination. Four algorithms – parkinsonism and the three OA algorithms – are entirely based on self-reported data collected via questionnaire. This is important because diagnosis of parkinsonism ordinarily involves a neurologist or movement disorder specialist, and diagnosis of musculoskeletal conditions often involves X-rays. The IHD algorithm may be used without ECG and results are better with self-reported data alone. Although the diabetes and CAO algorithms include blood or spirometry testing, many population-based studies collect blood samples as a matter of course. Also, spirometry is such a crucial pulmonary function test that it would be included in most studies of respiratory conditions. The validation study is noteworthy because we had no compliance issues as a result of asking control participants to undergo these tests.

Our estimates of sensitivity and specificity should be interpreted with spectrum effects in mind (Mulherin & Miller, 2002). Since we validated the algorithms in a population of definite cases and definite controls, the sensitivities and specificities shown in Tables 2a–e might overestimate algorithm performance in the CLSA. The CLSA will likely include a reasonable proportion of participants with only mild cases of disease that may be harder to detect using algorithms than the definite cases in our study.

Participant ages in the validation study were skewed towards the middle and upper-middle age range for eligible CLSA participants (i.e., 55 to 75 years of age). The CLSA will include participants above and below this range. To the extent that younger participants are likely to be healthier than participants in the 55- to 75-year age range, the presence of symptoms unrelated to disease in younger participants could lead to higher instances of false positives than our estimates of specificity would suggest. The same situation could occur in the upper age stratum (i.e., > 75 years). Compared to other age groups, the oldest CLSA participants might exhibit a greater number of co-morbidities and symptoms unrelated to any particular disease, thereby increasing the possibility of false positives.

Inaccurate estimation of specificity has implications for estimating disease prevalence in the CLSA, especially for rare conditions such as parkinsonism. For example, in a hypothetical sample of 1,000 persons, approximately 10 will have disease if the prevalence is 1 per cent. An algorithm with 100 per cent sensitivity and 90 per cent specificity will correctly identify all ten cases and incorrectly classify 99 individuals without disease as suffering from parkinsonism. For uncommon diseases, inaccurate estimation of specificity may translate into large variations in prevalence estimates.

Although the seven algorithms included in this pilot study will not correctly detect the true disease status of every participant, the low probability of false negatives and false positives is an acceptable trade-off, at least for common diseases, considering the richness of data that we will collect in the CLSA. Indeed, the alternative to these algorithms would be to forego data collection for important chronic diseases. Clinical diagnostic criteria, which require physician examinations, diagnostic testing, and physicians' interpretations of test results, would be difficult to standardize in a large study such as the 30,000-strong CLSA Comprehensive.

Table 2c: Chronic airflow obstruction algorithm sensitivities and specificities

Including/excluding Medication Question	Reclassification[a]	Sensitivity (%) (95% Confidence Interval)	Specificity (%) (95% Confidence Interval)
CAO (including medication question[a])	Possible CAO = test positive	100 (72–100)	0 (0–16)
	Possible CAO = test negative	100 (72–100)	80 (58–92)
CAO (excluding medication question)	Possible CAO = test positive	100 (84–100)	25 (11–47)
	Possible CAO = test negative	65 (43–82)	90 (70–97)

[a] **10 cases, 20 controls: medication data unavailable for 10 cases**
CAO = chronic airflow obstruction

Table 2d: Osteoarthritis algorithms sensitivities and specificities

Algorithm	Reclassification	Sensitivity (%) (95% Confidence Interval)	Specificity (%) (95% Confidence Interval)
OA Hand[a]	Uncertain or possible = test positive	100 (84–100)	95 (85–100)
	Uncertain or possible = test negative	65 (44–86)	95 (85–100)
OA Hip (including range of motion question)[b]	Uncertain OA = test positive Probable OA = test positive	100 (81–100)	85 (64–95)
	Uncertain OA = test negative Probable OA = test negative	87.5 (64–97)	95 (76–99)
	Uncertain OA = test positive Probable OA = test negative	100 (81–100)	95 (76–99)
	Uncertain OA = test negative Probable OA = test positive	87.5 (64–97)	85 (64–95)
OA Hip (excluding range of motion question)[b]	Uncertain OA = test positive Probable OA = test positive	100 (81–100)	95 (76–99)
	Uncertain OA = test negative Probable OA = test negative	75 (51–90)	95 (76–99)
	Uncertain OA = test positive Probable OA = test negative	100 (81–100)	95 (76–99)
	Uncertain OA = test negative Probable OA = test positive	75 (51–90)	95 (76–99)
OA Knee	Uncertain or probable = test positive	100 (84–100)	95 (76–99)
	Uncertain or probable = test negative	65 (44–86)	95 (85–100)

[a] Mild or severe asymptomatic = test positive
[b] 16 cases, 20 controls: recruitment halted after 16 cases
OA = osteoarthritis

The CLSA will afford the opportunity to conduct ongoing validation of the seven algorithms included in this pilot study, as well as the other algorithms included in the CLSA. For example, CLSA investigators will contact a proportion of participants who screen positive for parkinsonism in the actual CLSA, as well as a sample of screen negatives, and send them for a complete neurological examination. This will allow for direct assessment of sensitivity, specificity, and positive and negative predictive value.

In conclusion, we validated seven chronic disease ascertainment algorithms for use in the CLSA. The seven algorithms demonstrated an ability to correctly

Table 2e: Ischemic heart disease algorithm sensitivities and specificities

Including/excluding ECG	Reclassification	Sensitivity (%) (95% Confidence Interval)	Specificity (%) (95% Confidence Interval)
IHD (including ECG)	Probable = test positive Uncertain = test positive	100 (84–100)	60 (39–78)
	Probable = test negative Uncertain = test negative	80 (58–92)	85 (64–95)
	Probable = test positive Uncertain = test negative	100 (84–100)	75 (53–89)
	Probable = test negative Uncertain = test positive	80 (58–92)	70 (48–85)
IHD (excluding ECG)	Probable = test positive Uncertain = test positive	100 (84–100)	70 (48–85)
	Probable = test negative Uncertain = test negative	80 (58–92)	95 (76–99)
	Probable = test positive Uncertain = test negative	100 (84–100)	85 (64–95)
	Probable = test negative Uncertain = test positive	80 (58–92)	80 (58–92)

ECG = electrocardiogram
IHD = ischemic heart disease

detect disease for all cases and rule out disease for most controls. The parkinsonism algorithm had 100 per cent sensitivity and 100 per cent specificity. Although we developed these algorithms for the CLSA, they may be useful to accurately ascertain chronic diseases in other research settings as well.

References

Canadian Diabetes Association Clinical Practice Guidelines Expert Committee. (2008). Canadian Diabetes Association 2008 clinical practice guidelines for the prevention and management of diabetes in Canada. *Canadian Journal of Diabetes*, *32*(1 Suppl), S1–S201.

Duarte, J., Claveria, L. E., Pedro-Cuesta, J., Sempere, A. P., Coria, F., Calne, B. (1995). Screening Parkinson's disease: A validated questionnaire of high specificity and sensitivity. *Movement Disorders*, *10*, 643–649.

Hankinson, J. L., Odencrantz, J. R., & Fedan, K. B. (1999). Spirometric reference values from a sample of the general U.S. population. *American Journal of Respiratory and Critical Care Medicine*, *159*, 179–187.

Kriegsman, D. M., Penninx, B. W., van Eijik, J. T., Boeke, A. J., & Deeg, D. J. (1996). Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *Journal of Clinical Epidemiology*, *49*, 1407–1417.

Mulherin, S. A., & Miller, W. C. (2002). Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Annals of Internal Medicine*, *137*, 598–602.

Raina, P. S., Wolfson, C., Kirkland, S. A., Griffith, L. E., Oremus, M., Patterson, C., et al. (2009a). The Canadian longitudinal study on aging (CLSA). *Canadian Journal on Aging*, *28*, 221–229.

Raina, P. S., Wolfson, C., Kirkland, S. A., Keshavarz, H., Griffith, L. E., Patterson, C., et al. (2009b). Ascertainment of chronic diseases in the Canadian longitudinal study on aging (CLSA), systematic review. *Canadian Journal on Aging*, *28*, 275–285.

Rose, G. A., McCartney, P., & Reid, D. D. (1977). Self-administration of a questionnaire on chest pain and intermittent claudication. *British Journal of Preventive and Social Medicine*, *31*, 42–48.