Identifying patients with hypercholesterolemia

More than one blood sample is needed

MARK SPEECHLEY, PHD
SUSAN MCNAIR, MD
ALANNA LEFFLEY, BA
MARTIN BASS, MD

OBJECTIVE To compare the use of one and two blood samples for diagnosing hypercholesterolemia

DESIGN A test-retest substudy conducted as part of a randomized control trial designed to compare the effectiveness of different counseling strategies for lowering serum cholesterol, dietary fat, and dietary cholesterol in patients with moderate hypercholesterolemia

SETTING Thirty urban family practices

PARTICIPANTS One hundred forty-two patients provided two blood samples for total cholesterol (TC) level determination at two different times (test results were being used as an eligibility criterion for enrolment in the main trial).

MAIN OUTCOME MEASURES Number of subjects correctly classified to cholesterol risk category (normal < 6.2 mmol/L; moderate 6.2 to 6.9 mmol/L; high > 6.9 mmol/L) on the basis of one TC value and on the average of two TC values

RESULTS Overall misclassification rate on initial TC level was 22.5%. Overall false-positive rate was 19.0%, but false-positive rate for those initially assigned to the high category was 50%. Overall false-negative rate was 3.5%. Misclassification rates did not differ statistically on the basis of age, sex, blood pressure, smoking status, family history of coronary heart disease, presence of diabetes, obesity, the laboratory used, or whether the patient had fasted before giving blood.

CONCLUSIONS Single TC levels are too unreliable for diagnostic purposes, even if the subjects fast before testing. Family physicians should base their treatment decisions on the average of two cholesterol readings taken at different times 1 to 8 weeks apart.

OBJECTIF Comparer l'utilité d'obtenir un ou deux prélèvements sanguins avant de poser un diagnostic d'hypercholestérolémie.

CONCEPTION Une sous-étude test-retest effectuée dans le cadre d'un essai randomisé contrôlé visant à comparer l'efficacité de différentes stratégies de counseling à réduire la cholestérolémie, le gras alimentaire et le cholestérol alimentaire chez des patients porteurs d'une hypercholestérolémie modérée.

CONTEXTE Trente cliniques de médecine familiale oeuvrant en milieu urbain.

PARTICIPANTS Cent quarante-deux patients ont fourni deux échantillons de sang prélevés à des moments différents et destinés à mesurer la cholestérolémie totale (CT). Les résultats obtenus furent utilisés comme critère d'admissibilité à l'étude principale.

PRINCIPALES MESURES DES RÉSULTATS Nombre de sujets classés correctement dans l'une ou l'autre des catégories de risque d'hypercholestérolémie (normal < 6,2 mmol/L; modéré 6,2 à 6,9 mmol/L; élevé > 6,9 mmol/L) en se fiant soit à une seule valeur de la CT, soit à la moyenne de deux valeurs de la CT.

RÉSULTATS Le taux global de classification erronée sur la foi d'une seule donnée de TC fut de 22,5%. Le taux global de résultats faussement positifs fut de 19,0% mais le taux de faux-positifs chez ceux qui avaient été initialement assignés à la catégorie de risque élevé fut de 50%. Le taux global de résultats faussement négatifs fut de 3,5%. Les taux de classification erronée ne montrent pas de différence statistique lorsqu'on tient compte de l'âge, du sexe, de la tension artérielle, du tabagisme, des antécédents familiaux de coronaropathie, de la présence de diabète, d'obésité, du laboratoire utilisé ou du fait que le patient ait été à jeun ou non avant le prélèvement.

CONCLUSIONS La fiabilité d'un seul résultat de cholestérolémie totale est trop faible pour être utile à des fins diagnostiques même si les sujets sont à jeun avant le prélèvement. Les médecins de famille devraient baser leurs décisions thérapeutiques sur la moyenne de deux prélèvements effectués à des intervalles de une à huit semaines.

Can Fam Physician 1995;41:240-245.

Dr Speechley *teaches in the Departments of Physical Therapy, Faculty of Applied Health Sciences, and Epidemiology and Biostatistics, Faculty of Medicine, at the University of Western Ontario in London.* **Dr McNair** *and* **Dr Bass** *teach in the Department of Family Medicine in the Faculty of Medicine at the University of Western Ontario.*
Ms Leffley *is on staff at the Thames Valley Family Practice Research Unit in London.*

SERUM TOTAL CHOLESTEROL (TC) varies considerably in individuals over time.[1] As with blood pressure, an individual's cholesterol level can rise and fall substantially from day to day. Changes are large enough to move individuals in and out of categories labeled "normal" and "elevated" for clinical purposes. This basic biologic variability seems to have at least as much effect on TC levels as dietary change or procedural differences within and between laboratories.[1-3] For these reasons, estimating true TC levels for screening or for determining treatment requires at least two blood tests on separate occasions.

Knowledge of TC variability has been incorporated into programs for secondary prevention of coronary heart disease. For example, primary care guidelines have been published to assist physicians in assigning individuals to risk levels that are then used to help make treatment decisions. In Ontario, the Task Force on the Use and Provision of Medical Services recommended procedures endorsed by the Ontario Medical Association and the Ontario Ministry of Health.[4] These procedures include an initial TC test followed by another test if initial results are higher than 6.2 mmol/L. If initial values are lower than 6.2 mmol/L, no further tests are required, except for borderline values. If values on the two tests are widely discrepant, further testing is needed.

It is easy to envision instances in practice when two blood samples are not available and a hasty generalization of elevated cholesterol is made on the basis of a single sample. Patients are sometimes too busy to have a second test or unable to leave their work without financial penalty; patients in remote areas must travel long distances for blood testing; and for some patients, one set of test results is misplaced or spoiled.

In 1991 and 1992, we held several informal discussion sessions with family physicians as part of a dietary cholesterol reduction trial. During these discussions some physicians expressed frustration at the lack of specific details and the potential for misinterpretation or misapplication of the guidelines. Some indicated that they had been basing initial decisions about treatment on one TC value, believing this to be appropriate if taken from a fasting patient. This paper aims to demonstrate empirically the importance of observing the Task Force's basic requirement for two tests.

METHODS

Patient sample

Data for this paper came from 142 pairs of TC values collected as part of a randomized parallel-group trial of dietary cholesterol reduction[5] from 66 men and 76 women with a mean age of 52 years (SD = 12.6).

The cholesterol screening protocol used was from the then-new Task Force recommendations.[4] In applying these guidelines for research purposes, physicians were asked to identify from their charts patients in one of the risk categories (eg, man aged 35 to 60 years with one or more additional risk factors for coronary heart disease, anyone aged 20 to 69 with two or more risk factors, or anyone requesting screening). These patients were then asked to have a TC test. If first test results were > 6.2 mmol/L, patients were asked to give a second blood sample. We used first and second TC levels in our study. All samples were sent to one of three accredited laboratories in the city. The laboratories were aware that our study was under way as part of the larger trial but did not know which physicians or patients were involved.

Statistical methods

With parameters that display random variability over time within individuals, the theoretically optimal estimate of the attribute is the average of multiple measures. We took the average of

two TC measurements as the best single estimate of true individual TC level. We then used cross-tabulation to assess the error rate.

We used initial TC values to divide subjects into three groups: normal, moderately elevated, and high, just as we would in practice if all recommendations except repeat testing were followed. The rate of casefinding error was then calculated using the average of the initial and repeat measures.

Table 1. **Assignment to risk groups using initial TC test only versus using average of initial and repeat tests (N = 142):** *Normal results were below 6.2 mmol/L; moderate results were between 6.2 and 6.9 mmol/L; high results were above 6.9 mmol/L.*

		AVERAGE OF TWO TESTS		
RISK CATEGORY	INITIAL TEST	NORMAL	MODERATE	HIGH
Normal	5	5	0	0
Moderate	105	11*	89	5†
High	32	0	16‡	16

* *False positive (initially moderate category) = 11/105 = 10.5%.*
† *False negative (initially moderate category) = 5/105 = 4.8%.*
‡ *False positive (initially high category) = 16/32 = 50%.*

To support our assumption that the average of two measures is more reliable than either measure alone, we compared the coefficient of variation (CV) of each of the initial and repeat TC values with the averaged TC values (CV is the standard deviation divided by the mean). Because it standardizes the measure of variability (ie, SD) according to the magnitude of the original measure (ie, the mean), the CV allows direct comparison of the variability of two sets of values that have different means.

RESULTS

After 8 months, 25 physicians had screened 142 patients for the project. Initial mean TC score was 6.72 (SD = 0.42); repeat score was 6.47 (SD = 0.53).

Table 1 shows risk group assignment using initial results compared with the average of initial and repeat readings. The overall error rate was 32 (22.5%) of 142 (95% CI, 15.6% to 29.3%). The overall false-positive rate (subjects reassigned to any lower category) was 11 + 16 (19%) of 142 (95% CI, 12.5% to 21.5%). The false-positive rate for the 32 patients assigned to the high group was 16 (50%) of 32 (95% CI, 32.7% to 67.3%). Only five patients were reassigned to a higher risk category on the basis of averaged values for an overall false-negative rate of five (3.5%) of 142 (95% CI, 0% to 10%).

To see whether error rates differed systematically across levels of possible explanatory variables, we repeated the basic misclassification analysis separately for sex, age group, laboratory used, whether the subject had fasted or not before the test, and five coronary heart disease risk factors. Due to small cell sizes, the type of error (false-positive or false-negative) and the initial classification (moderate or high) was ignored, and all classifications were rated "correct" or "errors." The only statistically significant difference seen was the higher misclassification rate among normotensive patients ($P < 0.03$). If this finding is adjusted for the repeat statistical testing using Bonferroni's procedure (adjusted critical value: $\alpha/k = .05/9 = .006$), it becomes statistically non-significant.

DISCUSSION

The low reliability of TC measures has been known since the early 1950s.[1] Low reliability can be caused by instrumentation error (which encompasses all laboratory-related procedures including measurement technology used and practices followed for phlebotomy and for storing and handling blood); by the influence of dietary factors on TC; and by inherent biologic variability.

Instrumentation

The effect of instrumentation has been estimated by analyzing split samples of

the same sera. A recent Canadian risk factor prevalence study found the laboratory CV to be 1.3%.[2] Another recent study of 10 laboratories in the Little Rock, Ark, area found an overall CV of 3.2%.[3] National data from the United States suggest an average CV due to laboratory differences as high as 6.2%, but these data are older, predate recent standardization attempts, and might reflect earlier technologies.[6]

Diet

Researchers have tried to assess dietary influences on TC by analyzing serial TC measures from humans and other primates receiving controlled diets. With various controlled diets, human TC has ranged from 3.1% for obese patients on a formula diet to 9.0% for hospitalized schizophrenic patients.[1]

The range in CV is remarkable, particularly when compared with results of studies in which little or no dietary control was exerted (5.2% to 9.2%).[1] The CV calculated on our total sample (6.1%) is within the range of CV found in other studies and is, in fact, lower than that reported in studies in which little or no dietary control was instituted. The large overlap in ranges of CV between those on and not on controlled diets suggests that inherent biologic variability has at least as much effect on TC values over time as dietary intake.

At our orientation sessions, some physicians said they always requested that samples be given after fasting; others advocated fasting samples and fractionated lipids as a first screening step. Some reported they based decisions about therapy on the results of a single TC test, but added that they did this only if the sample had been taken after fasting. Belief in the importance of obtaining TC samples from fasting subjects might come from early laboratory literature stating that "5 ml of serum obtained after a 14-hour fast by the patient are required for a complete lipid analysis (cholesterol, triglycerides, and lipoprotein electrophoresis),"[7] or might be due to physicians' earlier training or to widely available and current general medical texts.[8] The task force guidelines[4] make no mention of fasting for TC or for fractionated lipids. More recently, Canadian practice guidelines explicitly advocate nonfasting samples for determining initial and repeat TC levels.[9]

However, the lowest CVs for initial and repeat samples came from those who reported that both samples had been taken after fasting. While we saw no significant difference in error rates between fasting and nonfasting subjects, this might reflect a lack of statistical power. A post-hoc power calculation reveals only about a 12% probability of detecting a difference in proportions this small between two groups of these sizes ($\alpha = .05$, two-tailed).[10]

Biologic variability

Biologic variability is addressed by repeat testing. The epidemiologic principle that screening is worthwhile only if an acceptable and efficacious treatment for the condition exists and if early detection confers a prognostic advantage[11] can apply in practice only if the screening test is capable of accurately assigning patients to risk categories. Our findings and the literature suggest that inherent biologic variability has a greater effect on accurate assignment to risk groups than laboratory procedures or patients' recent diets.

Task force recommendations

Task force recommendations address the need for repeat testing but neither explain the need nor emphasize the importance. A companion pamphlet aimed at patients was published late in 1991 by the Ontario Medical Association and the Ontario Ministry of Health.[12] In it, patients were told that "high or unexpected test results *will generally* be repeated" (italics added) before treatment decisions are made. The words "will generally" imply that in some cases tests will not be repeated, which, in view of the

many false-positive results from a single test, is unacceptable. Our findings suggest that between a third and two thirds of those who get a high reading will in fact get only a moderate reading on a repeat test, and more than 15% will be reassigned to a lower category.

One study found that, on retesting, 40% of the subjects had changed risk categories.[13] We found in our study that more misclassifications (28%) occur in the higher risk categories, which confirms the conclusions of another study.[14] Grossman[15] concludes that several TC values should be obtained before attempting to evaluate any therapeutic course for a given patient. Given the increasing popularity of cholesterol screening, however, and especially of any suggestion that whole populations or subpopulations be screened, this seems a prohibitively expensive option.

Recent Canadian practice guidelines[9] acknowledge the low reliability issue but their recommendations could lead to higher testing costs. For example, the guidelines suggest that all men aged 30 to 59 years who present for any reason should have a nonfasting cholesterol test. Then, for those with an initial level higher than 6.2 mmol/L, a repeat test should be conducted within 1 to 8 weeks. If the average of these two test results is higher than 6.2 mmol/L, physicians should "request fasting lipoprotein analysis, if available, or remeasure total cholesterol level."[9] This casefinding algorithm does not state how many times TC should be retested, nor what should be done to prevent the inevitable accumulation of serum TC values from forming an infinite feedback loop. The authors of the guidelines state that TC testing is "widely available and inexpensive," but because of the size of the population potentially involved, the compounded expense of a cheap test could still represent a heavy cost to the health care system. Indeed, the trade-off between cost and accuracy needs to be considered carefully, particularly since we have no firm evidence of the cost-effectiveness of cholesterol reduction in preventing coronary heart disease.[16]

Thompson and Pocock[17] found that the most substantial reductions in misclassifications are achieved by taking one additional measure and using the average, as was done in our study. In addition, literature provided by private laboratories[18] attempts to balance increased costs against improvements in diagnostic accuracy, such as taking the average of two values as the treatment baseline if the two readings are within 0.80 mmol/L, and the average of three readings otherwise. More research is needed to determine the basis for the 0.80 mmol/L difference criterion as well as to discover the optimal retesting interval for reducing the CV due to inherent biologic variability.

Conclusion

We conclude that single serum cholesterol levels are too unreliable to be used for clinical purposes. We urge family physicians to base their treatment decisions on the average of two cholesterol levels taken at different times between 1 and 8 weeks apart and remind them that repeated testing beyond that represents a costly search for marginal, if any, increases in accuracy. We also suggest further research to determine the optimal signal criterion for discrepant results, as well as the optimal interval for repeat testing of serum TC values. ■

Acknowledgment

This research was funded by a grant from the Ontario Ministry of Health's Health Care Systems Research Program. Dr Speechley thanks the Ontario Ministry of Health for support from a Career Scientist Award while much of this work was conducted.

Correspondence to: *Dr M. Speechley, Department of Physical Therapy, Faculty of Applied Sciences, University of Western Ontario, London, ON N6G 1H1*

References

1. Hegstead DM, Nicolosi RJ. Individual variation in serum cholesterol levels. *Proc Natl Acad Sci U S A* 1987;84:6259-61.
2. Connelly PW, MacLean DR, Horlick L, O'Connor B, Petrosovits A, Little JA, et al. Plasma lipids and lipoprotein and the prevalence of risk for coronary heart disease in Canadian adults. *Can Med Assoc J* 1992; 146(11):1977-87.
3. Kahn RF, McCord RS. Interlaboratory variability in serum lipid measurements. *Fam Pract Res J* 1991;11:415-20.
4. Toronto Working Group on Cholesterol Policy. *Recommendations to physicians for detection and management of asymptomatic hypercholesterolemia in adults.* Toronto: Ontario Medical Association and Ontario Ministry of Health, 1989.
5. Bass MJ, Lemaire DC, Speechley M, McNair SM, Kulalilaka H. Evaluation of dietary counselling strategies for cholesterol reduction in family practice patients. In: North American Primary Care Research Group. *Final program and abstracts.* Toronto: North American Primary Care Research Group, 1994.
6. Current status of blood cholesterol measurement in clinical laboratories in the United States: a report from the Laboratory Standardization Panel of the National Cholesterol Education Program. *Clin Chem* 1988;34:193-201.
7. Blood lipids. In: *Topics in laboratory medicine.* No. 5. Rexdale, Ont: MDS Laboratories, 1971.
8. Berkow R, editor. *The Merck manual of diagnosis and therapy.* 16th ed. Rahway, NJ: Merck Research Laboratories, 1992.
9. Canadian Task Force on the Periodic Health Examination. 1993 update. 2. Lowering the blood total cholesterol level to prevent coronary heart disease. *Can Med Assoc J* 1993;148(4):521-38.
10. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd ed. Hillsdale, NJ: Erlbaum, 1988.
11. Mausner JS, Kramer S. *Epidemiology: an introductory text.* 2nd ed. Philadelphia: WB Saunders, 1985.
12. *Information for patients about cholesterol.* Toronto: Ontario Medical Association and Ontario Ministry of Health, 1991 (rev ed).
13. Jacobs DR, Anderson JT, Hannan P, Keys A, Blackburn H. Variability in individual serum cholesterol response to change in diet. *Arteriosclerosis* 1983; 3(4):349-56.
14. Mogadam M, Ahmed S, Mensch AH, Godwin ID. Within-person fluctuations of serum cholesterol and lipoproteins. *Arch Intern Med* 1990;150:1645-8.
15. Grossman CM. Serum cholesterol variations in individual patients. *Atherosclerosis* 1988;71:193-5.
16. Sox HC. Screening for lipid disorders under health system reform [editorial]. *N Engl J Med* 1993;328(17):1268-71.
17. Thompson SG, Pocock SJ. The variability of serum cholesterol measurements: implications for screening and monitoring. *J Clin Epidemiol* 1990; 43(8):783-9.
18. MDS Laboratories. Cholesterol and the risk of coronary artery disease. In: *Topics in laboratory medicine.* Clinical Biochemistry No. C55. Rexdale, Ont: MDS Laboratories, 1989.

• • •