Inappropriate Requesting of Glycated Hemoglobin (Hb A₁c) Is Widespread: Assessment of Prevalence, Impact of National Guidance, and Practice-to-Practice Variability

Owen J. Driskell,1 David Holland,2 Fahmy W. Hanna,3 Peter W. Jones,4 R. John Pemberton,5 Martin Tran,1 and Anthony A. Fryer1*

BACKGROUND: Estimates suggest that approximately 25% of requests for pathology tests are unnecessary. Even in diabetes, for which international guidance provides recommended testing frequency, considerable variability in requesting practice exists. Using the diabetes marker, Hb A₁c, we examined (a) the prevalence of under- and overrequesting, (b) the impact of international guidance on prevalence, and (c) practice-to-practice variability.

METHODS: We examined Hb A₁c requests (519,664 requests from 115,730 patients, January 2001 to March 2011) processed by the Clinical Biochemistry Department, University Hospital of North Staffordshire, and prevalence of requesting outside guidance from intervals between requests was calculated. Requests were classified as “appropriate,” “too soon,” or “too late.” We also assessed the effect of demographic factors and publication of guidance, along with between-practice variability, on prevalence.

RESULTS: Only 49% of requests conformed to guidance; 21% were too soon and 30% were too late. Underrequesting was more common in primary care, in female patients, in younger patients, and in patients with generally poorer control (all \( P < 0.001 \)); the reverse generally was true for overrequesting. Publication of guidance (e.g., American Diabetes Association, UK National Institute for Health and Clinical Excellence) had no significant impact on under- or overrequesting rates. Prevalence of inappropriate requests varied approximately 6-fold between general practices.

CONCLUSIONS: Although overrequesting was common, underrequesting was more prevalent, potentially affecting longer-term health outcomes. National guidance appears to be an ineffective approach to changing request behavior, supporting the need for a multisystem approach to reducing variability.

© 2012 American Association for Clinical Chemistry

Evidence-Based Medicine and Test Utilization

Healthcare budgets worldwide are facing increasing pressure to reduce costs while maintaining quality. Laboratory test utilization has not escaped this pressure, particularly since pathology investigations are involved in 70%–80% of all healthcare decisions affecting diagnosis or treatment and cost the UK National Health Service (NHS), for example, £2.5 billion (US$4 billion) per year (representing almost 4% of total NHS expenditure) (1). Furthermore, many laboratories face an average annual increase in workload of 8%–10%.

One area of increasing importance is the management of inappropriate test requests. Many reports estimated that approximately 25% of pathology tests are unnecessary (1, 2)—and there have been higher estimates in some studies (2, 3)—representing a huge waste of resources. The Carter Review, a UK Department of Health–commissioned review of pathology services in England, acknowledges that there is “probably a similar amount of underrequesting” (1). There is little published evidence on this, possibly reflecting the difficulty in capturing such data.

The large variability in rates of test requesting between general practitioners (4–7) and hospitals (8) also supports the view that inappropriate requesting is widespread. This is observed even in diseases such as...
diabetes mellitus (DM), for which the American Diabetes Association (ADA) (9), the UK National Institute for Health and Clinical Excellence (NICE) (10, 11), and the Canadian Diabetes Association (12), among others, provide guidance on testing frequency. Unlocking the key to this variation and identifying the prevalence and impact of inappropriate requesting, both under- and overrequesting, would have major implications for reducing patient inconvenience and allow targeting of finite health service resources to areas of known patient benefit (13).

To identify the magnitude of inappropriate requesting, we determined the prevalence of requesting outside guidance with glycohemoglobin (Hb A₁c) in patients with DM as a model. We also examined the impact of publication of UK and US guidance on this prevalence by using data on 519 664 Hb A₁c requests in 115 730 patients over a 10-year period. Using requests from primary care, we then elucidated the variability in prevalence between general practitioners.

**Materials and Methods**

**PATIENTS**

We extracted data on all Hb A₁c test requests (n = 520 273) between January 2001 and March 2011 from the LabCentre Laboratory Information Management System database (Clinisys Ltd.) at the Department of Clinical Biochemistry, University Hospital of North Staffordshire NHS Trust, UK. During this period, there was little evidence (from clinical details supplied with requests) that Hb A₁c was being used as a diagnostic tool locally.

The data collected comprised request date, Hb A₁c result, anonymized patient identifier, age, sex, and requester. Data were processed in Excel 2007 (Microsoft) to remove external quality assurance–associated tests and assign requests to each unique patient ID, thereby allowing longitudinal assessment of intervals between results within individual patients. This left a core set of 519 664 requests from 115 730 patients. As expected, during the study period, new patients joined while others left the main cohort. Hence, the 115 730 comprised the patients who were included during the study period.

**DEFINITIONS OF UNDER- AND OVERREQUESTING**

UK NICE guidance documents for type I and type II DM recommend Hb A₁c testing at 2- to 6-month intervals in patients with unstable DM, with a measurement made at an interval of <3 months being used as an indicator of direction of change rather than of a new steady state (10, 11). In those with stable DM on unchanging therapy, intervals of 6–12 months are recommended. This is reinforced in the NHS Clinical Knowl-
with Stata statistical software, version 8.0, by use of “gllam” and “xtmixed” with the “xt” commands. Models were fitted with 2 levels: requests nested within patients. Within the models, the intercept and slope were both allowed to vary at the patient level (implying a reasonable assumption of correlation between the results of laboratory tests within patients). However, no significant patient-level variation was observed in any of the models fitted. Therefore, we used logistic regression models (Stata “logistic” command) at the test level; these analyses are presented in Results. We therefore examined the relationship between prevalence of requests too soon or too late and age, sex, requester, and degree of glycemic control by fitting logistic regression models with the Stata statistical software to obtain odds ratios and significance values.

Second, we assessed changes in prevalence over time and the impact of national guidance on a month-by-month basis from January 2003 to March 2011 with the previous 2 years as a rolling run-in period. For example, January 2003 prevalence estimates were obtained from tests requested during that month. They were identified as too soon or too late with prior data from January 2001 to December 2002 as run-in period. Similarly, January 2010 prevalence data were calculated from tests requested in January 2010 with data from January 2008 to December 2009 as run-in period. This 2-year run-in period provided a consistent estimate of prevalence of tests requested too soon, to allow like-for-like comparison of changes in prevalence over time. Although this shorter run-in period appears to consistently underestimate (by approximately 6% compared with the plateau at 9 years) overall prevalence of tests requested too late (see online Supplemental Fig. 1), it does allow month-by-month assessment.

### Results

#### Prevalence of Tests Requested Outside Guidance

Examining the 2010 dataset of 65 610 requests revealed that 11 614 (17.7%) were requested too soon (overrequests) relative to guidance and 16 291 (24.8%) were requested too late (underrequests). For the purpose of comparison with previous data [e.g., Lyon et al. (16)], we then excluded the first request in each patient (assumed to be appropriate by default), leaving a total of 54 537 repeat requests. Hence, all patients for whom only 1 test was requested were excluded from subsequent analyses. This further reduces any potential impact of diagnostic, rather than monitoring, requests.

Of the repeat requests, 21.3% were too soon (overrequests) relative to guidance (Table 1), which, on the basis of an Hb A1c test cost of £3 (US$4.80), would equate to £34 842 (US$55 747) in laboratory costs alone. However, the proportion of repeat requests considered too late (underrequests) was 29.9%, equating to £46 242 (US$73 987) in laboratory costs.

Table 1 also shows the effect of patient and requester factors, individually, on prevalence of inappropriate requesting. Compared with secondary care (hospital requesters), primary care (general practitioners) requested more tests too late and fewer tests too soon (both \(P < 0.001\)). However, as primary care requests constituted 87.5% of the total number of repeat requests, this group was the largest absolute contributor to tests requested outside guidance (Table 1). Tests requested too soon were more common in patients with stable diabetic control than those with poorer control (\(P < 0.001\) (Table 1), whereas the reverse was true for tests requested too late. To illustrate the distribution of retest intervals, we then plotted the relative frequency of requesting at weekly intervals for the 2003–2011 data.

---

**Table 1. Prevalence of inappropriate repeat requesting for Hb A1c.**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Total, n</th>
<th>Too late, n (%)</th>
<th>Too soon, n (%)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>54 537</td>
<td>16 291 (29.9)</td>
<td>11 614 (21.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>47 698</td>
<td>14 489 (30.4)</td>
<td>9274 (19.4)</td>
<td>&lt;0.001</td>
<td>0.82 (0.77–0.87)</td>
<td>2330 (34.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Secondary care</td>
<td>6783</td>
<td>1783 (26.3)</td>
<td>25 706 (46.1)</td>
<td>&lt;0.001</td>
<td>1.21 (1.17–1.26)</td>
<td>5607 (21.8)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>25 706</td>
<td>7995 (31.1)</td>
<td>14 489 (30.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28 810</td>
<td>8290 (28.8)</td>
<td>9274 (19.4)</td>
<td>&lt;0.001</td>
<td>1.12 (1.08–1.16)</td>
<td>6003 (20.8)</td>
<td>0.005</td>
</tr>
<tr>
<td>Age &lt;40 years</td>
<td>3766</td>
<td>1860 (49.4)</td>
<td>14 489 (30.4)</td>
<td>&lt;0.001</td>
<td>0.41 (0.38–0.44)</td>
<td>775 (20.6)</td>
<td></td>
</tr>
<tr>
<td>Age ≥40 years</td>
<td>50 771</td>
<td>14 431 (28.4)</td>
<td>10 839 (21.4)</td>
<td>0.266</td>
<td>1.05 (0.97–1.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well controlled</td>
<td>29 521</td>
<td>8099 (27.4)</td>
<td>8043 (27.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorly controlled</td>
<td>25 016</td>
<td>8192 (32.8)</td>
<td>3571 (14.3)</td>
<td>&lt;0.001</td>
<td>1.29 (1.24–1.34)</td>
<td>1.44 (0.43–0.46)</td>
<td></td>
</tr>
</tbody>
</table>

*Significance (P), odds ratios (ORs), and 95% CIs were calculated by using logistic regression.
set. These data showed prominent peaks at 3, 6, and 12 months for requests in both well- and poorly controlled patients from general practitioners, but a less clear (and generally earlier) pattern for hospital requests (Fig. 1, A and B).

Proportions of tests requested too soon and too late were significantly more common in female than male patients, although the absolute differences were small (reflected in relative-frequency plots) (Fig. 1, C and D). In terms of age, the data were first examined in 5-year bands, and this demonstrated a skewed distribution toward older age, as expected from the expected proportion of type I and type II DM patients (data not shown), although the actual number of type I and II patients in each age group was not known. On the basis of this examination and the shape of the distribution, we chose to dichotomize patients into groups of $<40$ vs $\geq 40$ years of age. Tests requested too soon were similar in the 2 age groups; relative-frequency plots, however, showed marked changes in pattern of repeat requesting with the younger age group and no discrete peaks, in contrast with the over-40 age group (Fig. 1, E and F). The older age group had significantly fewer tests requested too late ($P < 0.001$).

Overall, the group with the highest proportion of tests requested too soon comprised men $\geq 40$ years with well-controlled DM requested from secondary care (41.6%). The group with the highest proportion of tests requested too late comprised women $<40$ years with poorly controlled DM requested from secondary care (54.9%).

**IMPACT OF NATIONAL GUIDANCE**

We assessed impact of guidance with the 2003–2011 data and the 2-year run-in period. Overall, the introduction of UK Diabetes National Service Frameworks, NICE guidance, the UK GP Quality Outcomes Frameworks, and ADA guidance had no effect on prevalence of requesting outside this guidance (Fig. 2).

**BETWEEN-PRACTICE VARIABILITY**

We next examined the variability in proportion of tests requested too soon and too late between individual GP practices with the 2010 data. In 2010, there was a approximately 6-fold range (6%–32%) between practices in the proportion of tests requested too soon (Fig. 3A). The proportion of tests requested too late also showed substantial variability (approximately 6-fold; 9%–54%) (Fig. 3B).

**Discussion**

This study shows that pathology test requesting outside retest interval guidance in patients with DM is widespread. Indeed, on the basis of this guidance, 51.2% of repeat requests were inappropriate. We also demonstrate the lack of impact of publication of guidance on requesting patterns and the wide variability in prevalence between individual GP practices.

Our data indicate that repeat tests requested too soon account for approximately 21% of all requests for the DM marker Hb A$_{1c}$. This prevalence is consistent with previous estimates of overrequesting in general (1–3) and in DM in particular (17–19). For the first time, we also assessed the prevalence of underrequesting. Our data indicate that tests requested too late constituted approximately 30% of requests, indicating the need for more tests. However, our analysis may significantly underestimate, at least in some patients, the number of additional tests required to conform to guidance. For example, a request made 3 years after the previous test in a patient with well-controlled DM is, in our analysis, defined as a test requested too late. But the number of requests required to conform to guidance (every 6–12 months) would actually be at least 3 requests. If this assessment of missed tests is applied to our data, the minimum number of such tests in 2010 would be 30 313, equating to £90 939 (US$145 502) in laboratory costs, significantly more than that derived from tests requested too soon.

This study suggests that the largest proportions of tests requested too soon are from secondary care, well-controlled patients, and older patients, whereas the reverse was true for those requested too late. The observed requesting pattern, at least from primary care, with relative-frequency plots is similar to that of a Canadian study by Lyon et al. (16) that demonstrated peaks of activity at 1, 3, and 6 months corresponding to Canadian guidance (12). Our peaks are more pronounced at 3, 6, and 12 months (with a possible shoulder at 1 month), possibly reflecting UK guidance, although initial Hb A$_{1c}$ value appears to have little impact on requesting pattern. Furthermore, the requesting pattern from secondary care, and in younger patients, appears to be more ad hoc (with no clear pattern), highlighting these areas as requiring particular attention. Interestingly, we also observed the weekly spikes in requesting intervals described by Lyon et al. (16), confirming the tendency for phlebotomy to be performed on the same day each week within GP practices (see online Supplemental Fig. 2). The relative-frequency plots also suggest that some patients attend for tests 1–2 weeks on either side of the guidance limits. When we reanalyzed the data to include a 2-week “grace period,” the proportions of tests requested too soon and too late were reduced slightly from 21.3% to 18.2% and from 29.9% to 27.0%, respectively. These data suggest that requesters, at least in general practice, prefer a guide target retest interval, rather than a recommended range for repeat testing.
Fig. 1. Relative-frequency plots showing the distribution of repeat request intervals in well-controlled (initial Hb A1c < 7.0%) and poorly controlled (initial Hb A1c ≥ 7.0%) patients: primary care (A), secondary care (B), males (C), females (D), 40 years and older (E), < 40 years old (F).

Continued on page 911
Inappropriate Pathology Test Requesting

Fig. 1. Continued.
Our findings on overrequesting are consistent with those of Salvagno et al. (17), who showed that repeat testing within 3 months was higher in inpatients (particularly with Hb A1c values of <7%) than outpatients (particularly with values of >7%), although it should be stressed that interpretation of the inpatient data with definitive cutoffs (Table 1) should be treated with caution, as illustrated by the relative-frequency plots. Similarly, Akan et al. (19) showed that overrequesting was more common in inpatients than outpatients and, in the latter group, in those with an Hb A1c <7%. Neither study examined underrequesting or the effect of age. It also appears that this inpatient effect is not restricted to Hb A1c (18). The excess in secondary care does not appear to reflect a higher prevalence of difficult-to-control patients (who may require more-frequent testing) than in primary care, as the proportion of patients with an Hb A1c ≥7% (≥53 mmol/mol) was similar in the 2 groups (approximately 44%). The lower-than-average prevalence of tests requested too soon in younger patients mirrored the high prevalence of tests requested too late in this group. Indeed, our data showed that, in younger patients who were poorly controlled (Hb A1c ≥7%), 42.2% of tests (786/1862) were too late. The reasons for this may reflect, in part, limitations in the data set, as students who move away to further their education may have follow-up tests performed elsewhere. However, factors such as access to services for younger people and an increased proportion of type I patients in this group, who may be more difficult to engage with management of their condition, may also be important, as supported by the relative-frequency plots. Rankin et al. (20) suggested that reasons for poor knowledge and understanding of the condition in people with type 1 DM included diagnosis at a young age, assumption of decision-making responsibility by parents, lack of engagement when feeling well, transitions in care, inconsistency in information provision, and lack of awareness of poor or incomplete knowledge. These factors would present healthcare professions with challenges in providing optimum testing in this group.

The finding that individual GP practices or larger regions display large variations in number of tests requested, even in well-recognized scenarios, has been suggested (4–8). Indeed, O’Kane et al. (7) have recently identified this in the context of Hb A1c. However, the variability in request appropriateness has not
Fig. 3. Variability in proportion of repeat tests requested too soon (A) and too late (B) among the 87 GP practices in North Staffordshire, with the 2010 data set.
been assessed previously. Although a range of factors, including training or specialist knowledge in some practices, variation in basic medical training, patient demographics, attitudes to risk and litigation, length of local knowledge, and so on, may account for some of the variability, the standardization of best practice across the healthcare community should remain a goal. Approaches to reducing this variability have been discussed elsewhere (2, 3, 13, 21) and are beyond the scope of this article. Availability of data on previous test date and result has been suggested as a tool to limit overrequesting (3, 13). In our local health economy, secondary care has had access to general practitioners’ results throughout the study period, but the converse was not always the case. In October 2009, the rollout of electronic requesting systems for general practitioners to allow their access to secondary care results began, albeit slowly. When we examined the effect of this rollout on proportion of inappropriate tests (by comparing early vs late implementers), we found surprisingly little impact (data not shown). Indeed, we are now examining ways to enhance the educational content of the electronic requesting system with a view to reducing overrequesting.

What our data clearly show is the overall lack of conformity to national guidance and the lack of effectiveness of guidance in influencing requesting behavior. In the UK, NICE guidance for both type 1 and type 2 DM suggest minimum and maximum retest intervals for Hb A\textsubscript{1c} (10, 11), although, like most guidance on this subject, this guidance is often ambiguous and not well evidenced. Local guidance on testing frequency, as expressed in the UK Quality and Outcomes Framework for general practitioners, is extremely limited, focusing only on ensuring that each practice maintains a register of DM patients who have had an Hb A\textsubscript{1c} test within the previous 15 months. This emphasizes the need for local incentives and feedback to underpin robust, evidence-based guidance.

Other possible reasons for underrequesting in some cases can be proposed: lack of access to primary care practices outside working hours, comorbidity in elderly patients affecting mobility, lack of awareness of guidance, and so on. Similarly, patients may have tests sooner than guidance because of patient pressure, lack of awareness of the capability of the test itself, and so forth (13, 21). Although some of these reasons may be justifiable, they can in no way account for the magnitude of nonconformity. Smellie et al. (22) indicated that, although guidance exists, it is often spread among a range of literature sources and directed at laboratory specialists rather than requesters. Hence, our study supports the view that inappropriate testing should be a major target for health services to reduce costs (direct and indirect) and improve quality. Importantly, as current national guidance appears ineffective, new approaches that encompass the entire local healthcare economy, use a multisystem approach, and involve both patients and requesters are required (3, 13, 21). Furthermore, laboratories need to engage with requesters in defining the minimum time interval between tests needed to identify a significant change/difference.

In addition to the limitations outlined above, our analysis may be affected by factors such as the recently advocated use of Hb A\textsubscript{1c} in diagnosis, bias due to a relatively small number of atypical patients influencing intervals, the limited lead-in period for the 2010 data set, and the difficulty in differentiating between type I and type II DM. To mitigate these factors, we have (a) focused on repeat tests to remove single requests and collected data before publication of the recent guidance on use of Hb A\textsubscript{1c} in diagnosis, (b) performed multilevel modeling to examine whether the observations were influenced by patient-specific factors, (c) provided data on the magnitude of the underestimate of tests requested too late (see online Supplemental Fig. 1), and (d) separated patients by age (although we accept this is, at best, a crude surrogate of DM type in the absence of more definitive data).

Although we recognize that the absolute prevalence of inappropriate testing estimated by our analysis will have its limitations, our data clearly show that testing outside guidance (both under- and overtesting) is a major problem. Given that there may be legitimate reasons for some of this, our study suggests that there is likely to be wide variation in protocols among general practitioners. Although the Carter Report is being used to focus on reducing costs associated with unnecessary overtesting, its acknowledgement that there is probably a similar amount of underrequesting (1) is borne out by our findings. The cost implications of this, at least in terms of Hb A\textsubscript{1c} testing, are likely to be substantial. Extrapolating our data to the UK national context would mean approximately £5 million (US$8 million) in additional test costs alone for this 1 test. However, these costs may be offset by the potential benefits to patients and overall healthcare budgets arising from better control and reduced complications. Therefore, a more holistic approach to assessing the impact of pathology testing, taking account of the whole patient pathway, may be appropriate. Finally, it should be noted that our data are based on national retest interval guidance. The evidence base used to support these intervals is severely limited, and the potential impact on clinical outcome or patient experience of requesting patterns is currently poorly understood. Recent data from Fu et al. (23) suggest, as expected, that testing frequency is inversely associated with diabetic control. This study suggested that the optimal testing frequency

914  Clinical Chemistry 58:5 (2012)
to achieve an HbA1c below a target of 7% (53 mmol/mol) was 4 times per year. Turchin et al. (24) also showed that frequent HbA1c testing in DM patients resulted in shorter times to target HbA1c values, independently of confounders such as initial HbA1c value, treatment-associated factors, frequency of encounters with healthcare professionals, and patient demographics. More such data, to provide a clearer evidence base for the intervals suggested in national guidance, would be welcome.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

References


Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: D. Holland, National Pathology Benchmarking Service, Keele University; P.W. Jones, Keele University.
Consultant or Advisory Role: A.A. Fryer, National Pathology Benchmarking Service.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: O.J. Driskell, National Institute for Health Research Scientist fellowship award (HCS/08/011), supervised by A.A. Fryer.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We are grateful to the Health Foundation for the Shine award to A.A. Fryer and O.J. Driskell.

Clinical Chemistry 58:5 (2012) 915