

Online Supplementary Material

Verstappen W, van der Weijden T, Dubois WI, et al. Improving test ordering in primary care: the added value of a small group quality improvement strategy over classic feedback only a multicenter randomized trial. *Ann Fam Med*. 2004;2:569-575.

<http://www.annfammed.org/cgi/content/full/2/6/569/DC1>

---

**Supplemental Appendix. Upper Abdominal Complaints**

Primary care physicians received the feedback report on tests ordered for upper abdominal complaints 2 weeks before the small-group quality improvement meeting, together with the evidence-based guidelines on upper abdominal complaints. These guidelines recommend that there is no reason to order liver function tests for nonspecific upper abdominal complaints without jaundice. The risk of false-positive results is too large, because of the low prevalence of patients with liver diseases in primary care (4 to 5/1,000 patients). If physicians think screening is necessary, they are advised to order alanine aminotransferase (ALT) and  $\gamma$ -glutamyl transferase measurements for patients who do not have jaundice, and to order total bilirubin, ALT, and  $\gamma$ -glutamyltransferase measurements for patients who have jaundice. In short, there is never an indication to order more than 2 liver function tests in patients with upper abdominal complaints who do not have jaundice. Accordingly, aspartate aminotransferase (AST), lactate dehydrogenase, amylase, bilirubin, and alkaline phosphatase are considered to be inappropriate for patients with nonspecific upper abdominal complaints

At the meetings, physicians discussed their reports, compared their results with each other's and with the guidelines, and also discussed Bayesian decision rules to help them understand the probability of false-positive results in low-prevalence disorders. Another important topic of debate was the frequent requests by patients with nonspecific upper abdominal complaints to have blood tests. It took quite some effort and discussion to convince the physicians they had to change their routine for these cases. The next step was to try to implement the guidelines. Many physicians made plans for changes, such as "I will order fewer liver function tests, because I understand that these tests do not add useful information to what I know." Some local practice groups stated that they would use the same information brochure about nonspecific upper abdominal complaints.

The intervention arm physicians ordered on average 24 fewer tests per physician per half-year for upper abdominal complaints compared with the feedback physicians ( $P = .003$ ). For the intervention arm physicians, the number of inappropriate tests for this clinical problem decreased from 55 (SD 60) to 39 (SD 32), whereas in the feedback arm the number decreased from 60 (SD 63) to 56 (SD 54), meaning that the physicians in the nonintervention arm ordered 13 fewer inappropriate tests than the did feedback physicians ( $P = .002$ ) (Table 3).