SCREENING AND BRIEF INTERVENTION TARGETING RISKY DRINKERS IN DANISH GENERAL PRACTICE—A PRAGMATIC CONTROLLED TRIAL

ANDERS BEICH1*, DORTE GANNIK1, HENRIK SAELAN2 and THORKIL THORSEN1

¹Research Unit and Department for General Practice, Centre for Health and Community, University of Copenhagen, ²Medical Office of Health, City of Copenhagen, Denmark

(Received 4 October 2006; in revised form 20 May 2007; accepted 18 July 2007; advance access publication 11 September 2007)

Abstract — **Aims:** Recommendations for routine alcohol screening and brief counselling intervention in primary health care rest on results from intervention efficacy studies. By conducting a pragmatic controlled trial (PCT), we aimed at evaluating the effectiveness of the WHO recommendations for screening and brief intervention (SBI) in general practice. **Methods:** A randomized PCT (brief counselling intervention vs no intervention) involving 39 Danish general practitioners (GPs). Systematic screening of 6897 adults led to inclusion of 906 risky drinkers, and research follow-up on 537 of these after 12–14 months. Outcome measures focused on patients' acceptance of screening and intervention and their self-reported alcohol consumption. **Results:** Patient acceptance of screening and intervention —10.3% (N = 794) of the target population (N = 7,691) explicitly refused screening. All intervention group subjects (N = 442) were exposed to an instant brief counselling session while only 17.9% of them (79/442) attended a follow-up consultation that was offered by their GP. Consumption Changes At one-year follow-up, average weekly consumption had increased by 0.7 drinks in both comparison groups. As secondary findings, we observed an indiscriminate absolute risk reduction (ARR = 0.08 (95% CI: -0.02; 0.18)) in male binge drinking, but adverse intervention effects for women on the secondary outcomes (binge drinking ARR = -0.30 (95% CI: -0.47; -0.09)). **Conclusions:** The results of brief interventions in everyday general practice performed on the basis of systematic questionnaire screening may fall short of theoretical expectations. When applied to non-selected groups in everyday general practice SBI may have little effect and engender diverse outcome. Women may be more susceptible to defensive reactions than men.

INTRODUCTION

Primary health care is considered an ideal locus for health promotion and disease prevention. Routine screening in primary care for risky (harmful or hazardous) drinking and brief intervention for individuals screening positive in the form of feedback, information, and advice is widely recommended (Babor and Higgins-Biddle, 2000; Heather, 2002; US Preventive Task Force, 2004; Saitz, 2005; World Health Organization, 2006).

Brief intervention demonstrated efficacy (Moyer *et al.*, 2002) in research designs where two- or three-step assessment procedures secured homogeneous, compliant, and researchable study groups, thereby yielding higher internal study validity at the expense of external validity (Edwards and Rollnick, 1997; Beich *et al.*, 2003). Health promotion programs will usually only produce results (effectiveness) where efficacious programs target subjects in a manner acceptable to them (Flay, 1986).

Although brief assessment tools with fair sensitivity and specificity do exist, the actual impact of intervention towards a heterogeneous spectrum of risky drinkers simply identified through systematic screening remains untested in a naturalistic general practice setting (Beich et al., 2003; Whitlock, 2003). This evidence gap between efficacy (can it work?) and effectiveness (will it work?) has been recognized by some of the pioneers of brief intervention (Heather and Wallace, 2003). Pragmatic clinical trials designed to meet the needs of practitioners and decision makers could fill such gaps in evidence bases (Tunis et al., 2003).

*Author to whom correspondence should be addressed at: Research Unit and Department for General Practice, University of Copenhagen, Centre for Health and Community, Oster Farimagsgade 5, DK-1014 Copenhagen K, Denmark. Tel: +45 3532 7171; Fax: +45 3537 1282; E-mail: ab@gpract.ku.dk

In 1997, we initiated a study covering the grounding, validity, effectiveness, suitability and compatibility of a screening-based brief intervention approach recommended by the WHO (Babor *et al.*, 1992; Babor and Higgins-Biddle, 2001; WHO, 2006). Other results from our study that questioned the viability of systematic screening for risky drinking (Beich *et al.*, 2002) and the external validity of the evidence so far (Beich *et al.*, 2003), gave rise to a controversy so intense and emotionally charged that it called for a time-out in the publication process.

The aim of the study presented here was to qualify the evidence base of screening and brief intervention (SBI) by testing the recommendations. Our main objective was to ascertain to which extent brief intervention would trigger clinically significant and sustained consumption reductions among risky drinkers identified through systematic self-report questionnaire screening (Saunders and Aasland, 1987; Babor *et al.*, 1992).

METHODS AND MATERIAL

Protocol

The trial was testing the WHO SBI recommendations for primary health care that have not changed substantially since the present study was initiated in 1997. (Babor *et al.*, 1992; WHO, 2006). Systematic screening by use of the 10-tiem questionnaire, alcohol use disorders identification test (AUDIT) (Babor *et al.*, 1992; Babor *et al.*, 2001) administered either as an oral interview or as a self-report questionnaire is recommended. The screening total score is considered a simple way to provide each patient with an appropriate intervention, based on the level of risk. Brief intervention strategies are recommended for non-dependent risky drinkers (Babor and Higgins-Biddle, 2001).

The SBI methods were copied from the WHO collaborative study on brief interventions (WHO, 2001), but our study was conducted independent of the WHO study. The trial was approved by the Danish research ethics committee.

Setting. Denmark operates a health care system, where patients have access to free primary care from the general practitioner with whom they are listed. We invited 426 general practitioners (GPs) in four Danish counties individually to participate in a pragmatic trial. The written invitation made it clear that participation would at best be financially neutral and call for a close co-operation between GP and staff (typically an experienced nurse or assistant nurse with a central role in receiving the patients, giving advice, renewing prescriptions, deciding on the need for a consultation, and full access to patient files). Thirty-nine GPs finally volunteered to implement the program and participate in the study. Screening for lifestyle risk factors was not an established part of their job at the time, but they were willing to give SBI a try.

The doctors were given a full-day training course in the brief intervention elements by use of demonstration, role-play and reflective team methods. The elements of brief intervention were presented by using Feedback, Responsibility, Advice, Menu, Empathy, and Self-efficacy (FRAMES) (Miller and Sanchez, 1993), fortified with suggestions for handling resistance (Miller and Rollnick, 1991). The responsible staff members were trained individually and in groups in order to be able to include and exclude patients, introduce the program with consideration, and have the responsibility for all research procedures. They were paid an extra fee directly from the project for doing this work. During the screening period a telephone hotline was open for questions regarding screening and research procedures. This allowed doctors to focus exclusively on the intervention.

Eligible patients. Listed patients aged 18–64 years, scheduled to see their GP, and not fulfiling any exclusion criteria were invited to participate. They were carefully informed, ensured anonymity as regards any third parties, and asked to sign a brief informed consent. Those consenting (=participants) filled in an anonymous version of the AUDIT questionnaire (Babor et al., 1992) in the waiting room. We decided to have an upper age limit mainly for two reasons. First, we did not find the self-administered questionnaire suitable for elderly people (cognitive problems, level of cooperation, visual disorders). Second, the effect of age has not been systematically analysed as a possible influence on the AUDIT (Babor et al., 1992; Babor et al., 2001).

Screening exclusion criteria were: Severe acute illness; reading disability; speaking a foreign language, or illiteracy; mental or physical impairment; inebriation or currently being treated for an alcohol use disorder; pregnancy (for formal research ethical reasons).

Screening tool. The self-administered Danish AUDIT questionnaire was translated from English into Danish, compared to the original Norwegian (sister language) version of AUDIT (Saunders and Aasland, 1987), back-translated, reviewed by expert and layman panels, and tested on 600 patients in a pilot study before it was used in this study. AUDIT has previously been used in Danish psychiatric research (Nitschke et al., 1995), and in general practice research (Hansen et al.,

1999) but no full-scale validation has taken place. In accordance with the guidelines (Babor *et al.*, 1992; Babor *et al.*, 2001), a score of eight was chosen as the lower cut-off point.

Data collection. All patient data were questionnaire data. The screening questionnaire (AUDIT items, usual consumption in 12 g standard drinks of beer, wine, and spirits separately, together with demographic data) was administered pre-randomization, and was the primary source for baseline data collection. Post-randomization, a secondary anonymous baseline questionnaire was handed out to patients from both arms. They were instructed to complete it at home (recall diary of drinking last week and a few background questions) and mail it in a prepaid envelope directly to the project group.

Follow-up questionnaires (containing all screening questions as well as the ones from the secondary baseline questionnaire) were mailed to participants after one year followed by one reminder to non-responders. All data were entered into a database manually by double entry technique.

Assignment and final comparison groups. Participants completed the screening questionnaire in private, placed it in the envelope, and sealed it. The responsible staff member randomized all participants into one of two arms (*intervention arm* or *control arm*) by uncovering a symbol on the sealed envelope (like on a lottery scratch ticket).

Control arm patients dropped their envelope in a sealed ballot box at the reception, that is: they received no feedback on the result. Intervention arm patients took their envelope containing the completed AUDIT questionnaire to the surgery where the GP scored it. All patients scoring 8+ points were potential candidates for brief intervention. An alcohol dependency checklist based on the ICD-10 (WHO, 1993) was provided for GPs to run if the AUDIT score exceeded 12. Dependent patients should be referred for treatment and brief intervention conducted in the remaining cases.

Final comparison groups comprised patients who had not been referred for treatment and who stated a maximum weekly consumption of 35 drinks and scoring 8–21 AUDIT points (higher cut-off point suggested by Senft *et al.* (1997)) These upper thresholds were defined to avoid having responders in obvious need of alcoholism treatment in the comparison groups.

Masking. Blinding was not feasible, either for patients and GPs, or for outcome assessment and statistical analysis.

Intervention. The instant brief (10 min) intervention for patients who screened positive was based on the 'drink-less' protocol used by the WHO collaborative study on brief interventions (WHO, 2001). It should include feedback on present drinking, advice on reducing drinking with suggestions on how to do it, a self-help booklet, and an open invitation for a follow-up consultation at the earliest convenience. Doctors were asked to suggest an appointment and schedule one if appropriate. The intervention was thoroughly described in a doctor's manual and in brief on a flip chart for the doctor's desk

Sample size and outcome measures. Encouraging results from previous general practice trials (Wallace *et al.*, 1988; Anderson and Scott, 1992) made us aim for a mean betweengroup difference of five drinks (60 g) decrease per week in the usual consumption. This would demand a total of 250 subjects (control group reduction = 2, δ = 5, SD = 14, 2α = 0.05, β = 0.20). After having done a pilot study in which five

GPs had screened 600 patients, the GPs reported that some patients, particularly women, had reacted very defensively towards the screening or within the consultation room when a brief intervention was attempted (denial, anger, etc.). We anticipated gender differences due to this fact and we decided to simply double the sample size with the purpose of attempting to do a gender stratified effect analysis as a supplement to our hypothesis testing. We expected 10-15% to screen positive and a 60% response to the follow-up. We thus aimed for screening of 7000, receiving follow-up data from 500 patients. For the overall hypothesis, this would leave us with the trial power to detect a mean between group difference δ of three standard drinks.

In pragmatic trials, compliance is measured as an outcome. Measures reflecting level of patient acceptance and compliance are listed in Box 1.

Box 1. Measures reflecting patient program compliance

Availability and level of acceptance of the screening: The staff registered all subjects who

- Attended the surgery,
- Were eligible for screening,
- Accepted or refused the screening,

 Patient compliance with the intervention: To confirm
 that these activities had taken place, GPs registered
- Each brief intervention case accomplished (the first encounter)
- Subsequent clinical follow-up consultations accomplished

Drinking outcome measures are listed in Box 2.

Box 2. Effectiveness outcome measures

- Change in usual weekly consumption of beer, wine, and spirits separately as answered in the screening questionnaire (questions adopted from a Danish population study (Gronbaek et al., 1995))
 - Secondary consumption outcome measure
- Change in previous week's consumption of drinks, as per diary (Lemmens et al., 1992; Cohen and Vinson, 1995) completed in the secondary baseline questionnaire.

Secondary event outcome measures:

- Usual consumption from excess to within limits: men ≤ 21 drinks (252 g), women ≤ 14 drinks (168 g)
- Giving up binge drinking—six or more drinks per occasion on a weekly basis or more often
- AUDIT score decrease from 8+ to <8 points
- One of the above three events (not compensated by a reverse event in one of the other two). This composite measure was constructed post-hoc for enhanced sensitivity of the study.

Statistical analysis. All participants who underwent random allocation were primarily analysed according to group assignment (by 'intention-to-treat') as recommended for pragmatic trials (Roland and Torgerson, 1998; Macpherson 2004). We used a last-observation-carried-forward principle to account for missing data at follow-up. Changes in usual consumption were additionally analysed by protocol, that is, only patients who participated in the follow-up were included in this analysis.

Differences in mean consumption changes were analysed using two-tailed independent t-tests. Event outcome measures were calculated using absolute benefit increase (=ARR = absolute risk reduction): $treatment\ benefit$ (number with benefit per 100 interventions) and $screening\ benefit$ (number with benefit per 1000 screened = $treatment\ benefit$ × P_i , where P_i was the baseline prevalence of the particular risky drinking pattern in question). Confidence intervals were established using the Wilson method and CIA software (Altman $et\ al.$, 2000), whereas the Fisher Exact Probability Test was used to provide P-values.

Logistic regression analysis of the dichotomous dependent variable 'non-participation (participation) in research follow-up' was performed on the independent variables: group allocation; baseline consumption; AUDIT items and score; drinking concerns; change contemplations. Separate analyses were performed for men and women. The models were reduced by backward selection (a significance level of 0.05 was applied) and all the remaining covariates were allowed to have a different effect in the intervention and the control group. The procedure Proc Genmod in SAS was used for the analysis because it allows adjustment for a possible correlation among patients with the same practitioner.

Study flowchart

The initial screening was performed during eight consecutive working weeks in 1997. Research follow-up was conducted 12–14 months later. The 39 participating GPs were representative of Danish GPs regarding age (mean \pm SD = 48 \pm 5), gender (28% women), years in practice (mean \pm SD = 13 \pm 7), practice situation (25% rural practices), and organization (35% single-handed).

Participant flow and follow-up. During the screening period, 18–64-year-old visitors paid a total of 13 348 (personal contact) visits to the 39 participating GPs (Fig. 1). In 2482 of these visits a reason for exclusion was present, leaving 10 866 visits qualified for screening, of which 3175 were repeated visits during the study period. Thus, qualifying for screening were 7691 subjects, of whom 6897 accepted screening (participants) and 794 refused.

Participants (2600 men and 4297 women) were randomized to the intervention (3425) or control (3472) arms of the study. In all, 1087 (15.8%) screened positive (intervention 527; control 560). Alcohol dependency was indicated (AUDIT score >21, or, usual consumption >35 units/week) in 139 (2.0%), two intervention cases were missed by doctors, and 40 questionnaires had incomplete consumption data.

The study group thus comprised 906 patients, 442 in the intervention group and 464 in the control group. In all, 13.1% of participants (906/6897) were included in the study group: 23.3% of men (607/2600) and 7.0% of women (299/4297).

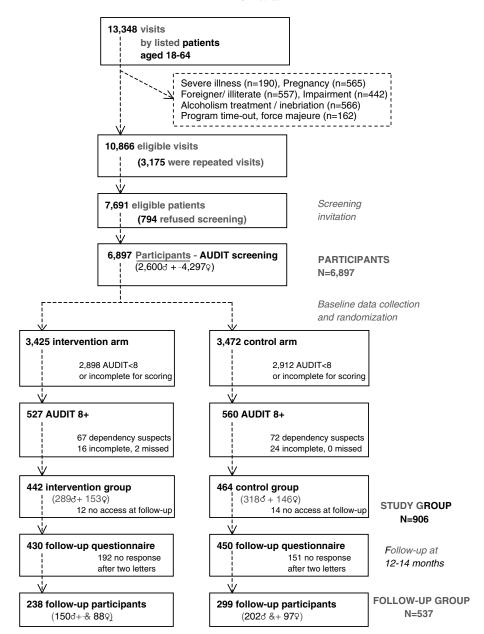


Fig. 1. Patients flowchart: Encounters, groups and losses.

The secondary baseline questionnaire was returned by 572 patients [intervention group 215/442 (48.6%), control group 357/464 (76.9%), P = 0.000000]. The one-year follow-up was completed by 537 patients [intervention group 238/442 (53.8%), control group 299/464 (64.4%), P = 0.001]. The overall follow-up rate was 61.0% in the group accessible for follow-up (N = 880, 26 had moved to an unknown address or died).

RESULTS

Baseline

Baseline characteristics. The study group comprised 906 AUDIT-positive drinkers. No differences were found between

the groups in terms of age, gender, cohabitation, and employment (Table 1).

In the study group, 24% of the men (147/607) and 17% of the women (52/299) stated a 'usual weekly consumption' above the threshold, while 38% (143/373) of the men and 32% (64/199) of the women who returned the secondary baseline questionnaire (N=572) reported consumption in the previous week as above the threshold (7-day recall diary). Weekly binge drinking (6+ units) was reported by 41% (246/607) of the men and 21% (63/299) of the women; the proportions for monthly binge drinking were 82% (497/607) and 69% (207/299), respectively.

Table 1. Baseline demographic and drinking characteristics of trial groups

Characteristic	Intervention	$n \text{ group}^* (N = 442)$	Control group* ($N = 464$)			
• Mean age ±SD, y	36.7	±12.8	36.3	±12.4		
• Men, N (%)	289	(65.4)	318	(68.5)		
• Living alone, N (%)	125	(28.3)	125	(26.9)		
• Employed or self-employed, N (%)	279	(63.1)	307	(66.2)		
\bullet Usual consumption, mean \pm SD, drinks	12.8	±8.7	12.9	±9.0		
• Usually above limit (women 14, men 21 drinks), <i>N</i> (%)	100	(22.6)	99	(21.3)		
$ullet$ Consumption last week, mean \pm SD, drinks ^a	17.0	±12.9	17.8	±14.2		
• Above limit last week, N (%) ^a	85/215	(39.5)	122/357	(34.2)		
• Binge drinking (6+ units) weekly+, N (%)	160	(36.2)	149	(32.1)		
• Binge drinking (6+ units) monthly+, N (%)	336	(76.0)	368	(79.3)		
• AUDIT-score 13–21, N (%)	104	(23.5)	105	(22.6)		

^{*} P-values for group differences were all above 0.10 (chi-square or t-test).

Programme compliance indicators

Patient acceptance of screening and intervention. A total of 10.3% (N=794) of the target population (N=7691) explicitly refused screening. Some of them were known or suspected by the GP or the staff to be heavy drinkers (Beich et al., 2002). Moreover, an unknown number claimed that they were unable to complete the questionnaire (e.g. forgotten spectacles, strained wrists, etc.) and were registered as such (impaired). After the first consultation, 17.9% (N=79) of the intervention group (N=442) returned for a follow-up consultation about drinking as suggested by their GP.

Effect analysis

Consumption changes. Both genders reported a modest (one drink or less) increase in their mean usual consumption from baseline to follow-up (Table 2—by intention-to-treat). No significant differences were found between the groups, and analysis by protocol did not change this result (Table 2—by protocol).

Patients returning the supplementary baseline questionnaire containing the one-week drinking diary (N=572) were included for a supplementary analysis of consumption changes (Table 3). In all, 49% (215/442) of the intervention group; 77% (357/464) of the control group had returned this questionnaire. Mean weekly consumption declined by 0.6 drinks among men in the intervention group, but rose by 0.8 drinks in the control group (P=0.31). Among women, it rose by 1.7 drinks in the former group, but declined by 0.1 drink in the latter (P=0.23).

Event outcome. At follow-up, 26% of the whole study group (233/906) had managed to moderate at least one risky drinking pattern (Table 4). The overall study group differences were small (\leq 2%), statistically insignificant, and not in favour of intervention. Intervention tended to have a positive influence on men (strongest for weekly binge drinking: ARR = 8% (95%CI: -2%;18%)), while women displayed

adverse intervention effect tendencies on all accounts. Female weekly binge drinking was affected by intervention in a negative way: ARR = -30%(95% CI: -47%; -9%), corresponding to P = 0.007.

Drop-out

Drop-out regression analysis (Table 5) showed that men having received intervention were lost to follow-up more often (odds ratio (OR) = 1.75, 95% CI: 1.23; 2.51) than men who had not, and a 10-year age increase was associated with an OR of non-participation of 0.85 (95% CI: 0.75; 0.98). For women, the likelihood of non-participation rose significantly with baseline consumption: a five-drinks/week increase corresponded to an estimated drop-out OR of 1.31 (95% CI: 1.10; 1.56). Intervention and age also tended to be predictors of drop-out for women, but these findings were statistically insignificant.

DISCUSSION

Key findings

We found no support for the main hypothesis that brief intervention would cause self-reported weekly consumption to decline among drinkers identified by general AUDIT screening. On the secondary outcome measures we found a trend towards a positive effect of intervention on male binge drinking, and an adverse outcome for women on all measures. Less than one in five intervention patients returned for a follow-up consultation offered as part of the brief intervention, thus disclosing a limited demand for further counselling by their GP on those conditions.

Limitations of the present study

The 39 participating GPs were likely to have a higher motivation than the usual one among GPs because they

^a 49% of the intervention group and 77% of the control group returned the weekly recall diary (supplementary baseline questionnaire).

Table 2. Changes in 'usual consumption' (12 g drinks) from baseline to follow-up in comparison groups—by intention-to-treat (ITT) and by protocol

							oj protoc	01							
	All							Men		Women					
	Interv	ention	Cor	ntrol	t-test	Interv	Intervention		ntrol	t-test	Intervention		n Control		t-test
		Usual consumption (units/week) (by ITT) ^b													
N participants		42 n (SD)		64 n (SD)			89 n (SD)		18 n (SD)		153 Mean (SD)		146 Mean (SD)		
Baseline 12 months Change	12.8 13.5 +0.7	(8.7) (11.1) (7.4)	12.9 13.6 +0.7	(9.0) (11.7) (8.1)	P = 0.77 P = 0.83 P = 0.99	14.9 15.6 +0.7	(9.1) (11.6) (8.1)	15.1 15.6 +0.5	(9.1) (11.9) (8.0)	P = 0.74 P = 0.96 P = 0.76	8.8 9.5 +0.7	(6.0) (8.1) (5.9)	8.2 9.3 +1.1	(6.7) (10.0) (8.3)	P = 0.41 P = 0.85 P = 0.63
Percent change	+5	+5.5% +5.4% +4.7% +3.3% +8.0% +13.4% Usual consumption (units/week) (by protocol) ^a													
N participants		224 288 144 an (SD) Mean (SD) Mean (SD)				99 n (SD)		80 Mean (SD)			89 n (SD)				
Baseline 12 months change Percent change	12.6 14.0 +1.4 +11	(8.6) (12.7) (10.4) 1.1%	12.8 14.0 +1.1 +8	(9.0) (13.0) (10.2) .6%	P = 0.76 P = 0.98 P = 0.77	14.9 16.3 +1.4 +9	(9.1) (13.6) (11.5) .4%	15.2 16.0 +0.8 +5	(9.1) (13.2) (10.1) .3 %	P = 0.76 P = 0.83 P = 0.60	8.5 9.8 +1.3 +15	(5.8) (9.5) (8.1) 5.3%		(6.1) (11.3) (10.6) 3.7%	P = 0.34 P = 0.80 P = 0.75

^a By protocol means that only those with complete follow-up data (N = 512) were included for analysis.

Table 3. Retrospective diary consumption changes from baseline to follow-up

	All				Men						Women				
	Interve	ention	Cor	ntrol	t-test	Interv	ention	Cor	ntrol	t-test	Interv	ention	Cor	ntrol	t-test
		Consumption last week (units/week)							s/week)						
N participants	21 Mean		-	57 n (SD)			134 Mean (SD)		39 n (SD)		81 Mean (SD)		118 Mean (SD)		
Baseline 12 months		(12.9) (13.2)	17.8 18.3	(, ,	P = 0.49 P = 0.46	20.1 19.7	(13.6) (13.7)		, ,	P = 0.85 $P = 0.37$	11.9 13.4	(9.8) (11.3)	12.5 12.3	(9.8) (12.4)	P = 0.71 $P = 0.30$
Change Percent change	+0.3 +1.7	(10.9) 7%		(12.2) .1%	P = 0.80		(11.4) .0%		(13.3) .9%	P = 0.31	+1.7 + 1 4	(10.1) 1.3%	-0.1 -0.1	(9.4) .8%	P = 0.23

Included for analysis were the sub-sample (N = 572 patients) who had returned the secondary baseline questionnaire that included a drinking diary. The follow-up rate was 72.0% (N = 412), missing data (N = 160) at follow-up were substituted by baseline values.

were self-selected. This might, in turn, influence the study outcomes by favouring positive results in our trial.

The screening methodology used here represents 'true' screening (Wilson and Jungner, 1968) as recommended (Saitz, 2005): it is applied to large groups irrespective of the reason for encounter or other situational factors. It should not be confused with a brief assessment of an individual patient's drinking, even if this is also often referred to as screening.

The trial attempts to assess the long-term effect of a onetime brief counselling session with an open invitation for follow-up consultations. Because there was too little demand in the intervention group we cannot conclude anything about repeated counselling sessions, and we do not have data to shed light on any possible short-term effects of the brief intervention. Although we asked GPs to run a dependency checklist if the patient scored 13+ on AUDIT, and we excluded subjects who stated a consumption of above 35 drinks/week, it could be argued that we might have included alcohol-dependent subjects in our effect analysis (AUDIT score 8–21). More than 75% of the study group had scored 8–12 on AUDIT, and subgroup analyses on subjects with AUDIT scores between 8 and 12, 8 and 13, 8 and 14, etc. were carried out. These analyses did not reveal better results, the group differences remained small and the tendencies for men and women were stable.

The possibility of a cross-over effect is present. Having been primed by the screening questionnaire, some control patients may have become more aware of their drinking and maybe more open to discussing alcohol use issues with the GP. But the early randomization of all participants

^b By ITT (intention-to-treat) means that all study group members (N = 906) were included for analysis (Last Observation Carried Forward, that is, missing data at follow-up were substituted by baseline values).

Table 4. Excessive drinking patterns at baseline and positive changes at follow-up—Analysis by intention-to-treat: Absolute benefit per 100 interventions (treatment benefit) and per 1000 patients screened (screening benefit)

AII					Men						Women						
Number screened			6897			2600						4297					
	Inter	vention	Control		Test	t Intervention		Control		Test	Intervention		Control		Test		
Participants, N	442		464			289		318			153		146				
					τ	Jsual	consun	ption above	weekly	limits							
Baseline, N	100		99			72		75			28		24				
Within limits at follow-up, N (%)	18	(18.0)	20	(20.2)		14	(19.4)	15	(20.0)		4	(14.3)	5	(20.8)			
treatment benefit (95% CI)		-2	(-13 to 9)		P = 0.73		-1	(-13 to 12)		P = 1.00		-7	(-28 to 14)		P = 0.72		
screening benefit (95% CI)		-1	(-4 to 3)				0	(-8 to 7)				-1	(-3 to 2)				
					1	Binge	drinki	ng (weekly o	r more	often)							
Baseline, N	160		149			132		114			28		35				
No binging at follow-up, N (%)	36	(22.5)	33	(22.1)		34	(25.8)	20	(17.5)		2	(7.1)	13	(37.1)			
treatment benefit (95% CI)		0	(-9 to 10)		P = 1.00		8	(-2 to 18)		P = 0.13			-30(−47 to −9))	P = 0.007		
screening benefit (95% CI)		0	(-4 to 4)				8	(-2 to 17)					-4(-7 to -1)				
							AUD	T score >8	points								
Baseline, N	442		464			289		318			153		146				
AUDIT ≤ 8 at follow-up, N (%)	88	(19.9)	99	(21.3)		51	(17.6)	58	(18.2)		37	(24.2)	41	(28.1)			
treatment benefit (95% CI)		-1	(-7 to 4)		P = 0.62		-1	(-7 to 6)		P = 0.92		-4	(-14 to 6)		P = 0.51		
screening benefit (95% CI)		-2	(-9 to 5)				-1	(-16 to 13)				-3	(-10 to 4)				
					(Cha	nging	g) at lea	st one risky	drinkir	ıg pattern	a						
Baseline, N	442		464			289		318			153		146				
One change at follow-up, N (%)	114	(25.8)	119	(25.6)		75	(26.0)	73	(23.0)		39	(25.5)	46	(31.5)			
treatment benefit (95% CI)		0	(-6 to 6)		P = 1.00		3	(-4 to 10)		P = 0.40		-6	(-16 to 4)		P = 0.30		
screening benefit (95% CI)		0	(-7 to 8)				7	(-9 to 23)				-4	(-11 to 3)				

^a Only counted positive if the event is not counterbalanced by a reverse change (negative development) in one of the other two.

made it possible for us to avoid that any feedback on the screening result was given to control group patients, in order to minimize this effect. Controls were merely taking part in a rolling survey, and neither the patient nor the GP knew about the screening result at any point.

Self-reported drinking is subject to some uncertainty (Dawson, 1998; Rehm, 1998; Room, 1998). The summary measure raises questions about accuracy, and the diary method ignores that a single subject's consumption often varies substantially over time. Although the last week consumption measure showed a higher consumption than the self-reported usual consumption at both baseline and follow-up, the results for usual consumption and consumption last week were comparable and the differences over time as well as the between

group differences were small. In order not to discourage GPs and patient participation and jeopardize our focus on every-day effectiveness, we refrained from collecting any collateral information.

Limited follow-up rates are not uncommon in pragmatic trials (Kendrick *et al.*, 1995; Peveler *et al.*, 2005). However, validity was reduced by the differential loss to follow-up, and no advanced model for replacing the missing data could make up for the potential biases introduced by these losses.

The challenge is, accordingly, to plausibly explain the direction of the potential biases introduced by the limitations of our study. We found no reason to expect a better outcome, in general, among those lost to follow-up than among participants. We used a 'last-observation-carried-forward principle',

Table 5. Baseline predictors of research drop-out (follow-up non-participation)

Gender	Variables	Effect of	OR	95% CI
Men	Allocation Age	Intervention +10 Years		
Women	Consumption $(t = 0)$ Allocation Age Consumption $(t = 0)$	+5 Drinks Intervention +10 Years +5 Drinks	1.35 0.91	0.95; 1.13 0.83; 2.12 0.74; 1.11 1.10; 1.56

Independent variables entered: Group allocation, baseline consumption, AUDIT score, the ten AUDIT items, drinking concerns, change contemplations.

Separate analyses were performed for men and women. The effect of age and consumption could be assumed to have the same effect in the intervention as well as the control group. The models were reduced by backwards selection based on a 5% significance level, and all the remaining covariates were allowed to have a different effect in the intervention and the control group.

In the Table, the gender-specific effect of the remaining three predictors is estimated.

assuming no change among those lost. However, heavier-drinking women and men having received the intervention were lost more often than other groups, and further aggravation of drinking problems may have taken place in some cases. We found no reason to assume that lost opportunities for registering positive changes in the intervention group would outnumber such opportunities in the control group. In fact, we believe that there are several conditions that might lead to an overestimation of possible positive effects (i.e. focus on positive changes, no blinding of participants, outcome measures rely on self-reports, etc.).

AUDIT defines a rather heterogeneous group of risky drinkers. Some were included because of immature binge drinking patterns, others because of negative consequences of drinking in the past. In spite of their positive self-selection, some GPs occasionally became uncertain about both, role legitimacy and role adequacy, when working in this way with these often surprisingly defensive drinkers (Beich et al., 2002). Agenda setting by screening is, by nature, confrontational, and confrontation is known to engender resistance (Miller et al., 1993). The question is whether brief intervention incorporating advice towards a highly defensive patient makes any sense at all.

Devoting a single day for training in how to perform brief interventions may be inadequate. On the other hand, one of the advantages of brief intervention is that you do not need much training to be able to do it (Babor *et al.*, 2001). Furthermore, four doctors had attended a three-day 'helping people change' course focused on drinking problems prior to the programme, and neither experienced less difficulty nor produced better results.

The low level of acceptance of the programme (participation in, compliance with, or adherence to (Flay, 1986)) is likely to have influenced its outcome. Still, it can be argued that pragmatic clinical trials measure effectiveness (the benefit that treatment produces in routine clinical practice (Roland and Torgerson, 1998)) and that lack of compliance is, indeed, an outcome when reasonable attempts have been made to encourage compliance (Godwin *et al.*, 2003).

Pragmatic trials should compare clinically relevant alternatives. Because screening was included as a premise in our study we might have compared two unfamiliar alternatives. Outcomes from SBI programs should ideally be compared to the outcomes of usual care. Usual care includes a large number of encounters in which drinking assessment and counselling may surface on the initiative of either, the doctor or the patient, in a continued doctor-patient dialogue. This dialogue is rooted in the doctor-patient relationship and typically shaped by the patient's lack of well-being, specific complaints, objective findings (maybe vague or inconclusive or none), or maybe the patient's (or his/her family's) general ability to function.

Possible mechanisms and explanations

A screening-based brief intervention applied in the real world context of general practice may not have the effect suggested in the efficacy studies and meta-analyses published so far. Some of these unselected patients identified by systematic screening were defensive (Beich *et al.*, 2002) and more than four out of five declined further consultations on the alcohol issue. The negative effects among women may reflect such defensiveness. Resistance or dissonance within the consultation is known to lower the likelihood of change (Miller *et al.*, 1993) and women may be more sensitive to criticism regarding drinking habits than men (Gomberg, 1988).

Comparison with present evidence

Most published brief intervention studies that did not find any additional benefit to brief intervention showed a decrease in alcohol use both in experimental and control groups. This parallel decrease was explained either by a Hawthorne effect, a natural history effect, or by regression to the mean. We did not find any decrease in alcohol use in either groups, which can be explained by the fact that we included 83% of all patients screening positive in the AUDIT, and that the intervention was provided without delay as recommended (Babor et al., 1992; Babor et al., 2001). This strategy differs from the two- or three-step screening models and special office visits used in previous brief intervention trials and have resulted in a less selected study group for which we can expect results to differ from studies on much more homogeneous and selected groups. Our results somehow confirm that over time alcohol problems may take a turn for the worse for some risky drinkers. A preventive effort towards risky drinkers is justified, and we should definitely not pin our faith on natural history or regression to do the job for us in all cases.

The WHO collaborative study on SBI developed and tested screening procedures (Saunders and Aasland, 1987) and brief interventions (Babor and Grant, 1992) separately, whereas the SBI recommendations that followed were not tested in a research set-up in real life conditions.

The most widely quoted study on brief interventions is probably a US efficacy study (Fleming *et al.*, 1997) that screened 17 695 primary care patients, and found 2925 (17%) with risky drinking habits, of whom 774 (4.4%) were included

after further assessment and interview, and 723 (4.1%) were followed up. For comparison, we screened only 6897 patients, found 1087 (15.8%) risky drinkers, of whom 906 (13.1%) non-dependent drinkers became study participants and 537 (7.8%) were followed up. Another primary health care trial (Senft *et al.*, 1997) used the AUDIT to define the trial group, and included 78% of those who screened positive. The main intervention was not delivered by the GP, but by a skilled counsellor. Nevertheless, only a modest, temporary reduction in the drinking frequency but not in the overall amount of alcohol consumption was achieved in this study.

Several systematic reviews have confirmed the efficacy of brief intervention in reducing risky levels of alcohol consumption in non-dependent individuals (Bien et al., 1993; Kahan et al., 1995; Wilk et al., 1997; Poikolainen, 1999; Moyer et al., 2002; Ballesteros et al., 2004). Until lately, little attention has been paid to the issue of the external validity of these results (Edwards and Rollnick, 1997; Beich et al., 2003). A recent systematic review (Bertholet et al., 2005) on the efficacy of brief alcohol intervention concluded that intervention could reduce alcohol consumption, but that the effectiveness of such interventions in daily practice remains unexplored. This gap between efficacy and effectiveness has also been recognized by some of the pioneers of brief intervention (Heather and Wallace, 2003). The proposals to overcome this gap have, so far, focused on the search for even better screening tools and better implementation strategies for SBI programs.

Recently, the WHO collaborative project on SBI implementation that inspired our work in the first place, published their final report (Heather, 2006). In this report of action research in which 'serious difficulties were encountered' in practically all participating countries it is concluded that screening is an imperative and implementation is still the challenge. Our papers (Beich et al., 2002, Beich et al., 2003) are once again claimed to have had 'some detrimental effects on efforts to promote screening for hazardous and harmful alcohol consumption in general medical practice'. The facts that we actually implemented screening and intervention in the practices of 39 highly motivated GPs, and that their screening experiences (not attitudes) were negative (Beich et al., 2002), are ignored in this report although it contains a separate chapter headed 'Denmark'.

Research and clinical implications

The lack of external validity from previous general practice trials (Beich et al., 2003), the results of the present pragmatic trial, and the experience of the participating GPs (Beich et al., 2002) point to the need for studying health behaviour interventions as complex interventions (Campbell et al., 2000). Should we strive to adjust everyday practice to the conditions under which the available evidence is obtained rather than strive to ensure that our proposals are suitable for everyday practice? It may, indeed, be necessary to devote more attention to the perspectives of the patient and the provider (theoretical and modelling phases) and to describe the constant and variable components of replicable and acceptable interventions (exploratory trials) before launching both, efficacy and effectiveness trials.

The prevention paradox that few patients personally benefit from systematically delivered preventive interventions is likely to embrace SBI for alcohol problems. The GP here faces a profound conflict of interests: to care for the health of the nation (and sometimes the ambitions of an anonymous third party) while trying to do their best for the personal health needs of the individual patient. If every instance of risky alcohol use (including binge drinkers) were to be addressed at every opportunity as suggested (Saitz, 2005), the GP would be consulting with around half of the adult population for ongoing alcohol education, advice, counselling or treatment. Family doctors would have to sacrifice other activities (Yarnall *et al.*, 2003) and to spend resources on screening activities of unknown effectiveness (Whitlock, 2003; Beich *et al.*, 2003).

If the GP seizes the right moment for addressing the issue, and does so in a manner that allows the patients to see advice as an integral part of the GP's care activities, drinking assessment and advice giving may be likely to help a larger fraction of drinkers and to provoke less dissonance and resistance than advice simply offered to everyone at risk, regardless of the patient's agenda, receptivity and state of mind. A recent paper (Sussman *et al.*, 2006) on clinicians' preventive counselling decisions describes a complex set of factors that influence their decisions to provide preventive counselling and suggests that we move beyond linear models of behaviour change to recognize the complex environment of primary care.

GPs should nevertheless have the skills necessary for raising sensitive lifestyle issues like drinking, smoking, diet, and exercise, whenever appropriate, in an atmosphere of consonance (Miller and Rollnick, 2002). Sufficient skill-building in health practitioner training, and continuing education to secure an ongoing and 'healthy' dialogue with the patient is an ongoing challenge of the future.

Acknowledgment — We are indebted to the 39 GPs and their receptionists for their great efforts with screening, intervention and data collection. We gratefully acknowledge the comments and help in preparing and carrying out the study by AlkoholGruppen, Charlotte Hindsberger, Poul A. Pedersen and Hanne Hollnagel at the Research Unit and Department of General Practice, University of Copenhagen, and the help by Claire Lane at Department of General Practice, Cardiff University, in our writing up this paper. This work was supported in part by grants from Alkoholpolitisk Kontaktudvalg (Danish Ministry and Board of Health), Forskningsfonden (Association of County Councils in Denmark), and Fonden til Laegevidenskabens Fremme.

REFERENCES

Altman, D. G., Machin, D., Bryant, T. N. et al. (2000) Statistics with Confidence. 2nd edn. BMJ Books: London.

Anderson, P. and Scott, E. (1992) The effect of general practitioners' advice to heavy drinking men. *British Journal of Addiction* 87, 801, 900

Babor, T. F., de la Fuente, J. R., Saunders, J. B. et al. (1992) The AUDIT: Guidelines for Use in Primary Care. WHO: Geneva.

Babor, T. F. and Grant, M. (1992) Project on Identification and Management of Alcohol-Related Problems. Report on Phase II: A Randomized Clinical Trial of Brief Interventions in Primary Health Care. WHO: Geneva.

Babor, T. F. and Higgins-Biddle, J. C. (2000) Alcohol screening and brief intervention: dissemination strategies for medical practice and public health. *Addiction* **95**, 677–686.

Babor, T. F. and Higgins-Biddle, J. C. (2001) *Brief Intervention for Hazardous and Harmful Drinking: A Manual for Use in Primary Care*. WHO: Geneva.

- Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. *et al.* (2001) *The AUDIT: Guidelines for Use in Primary Care*, 2nd edn. WHO: Geneva.
- Ballesteros, J., Gonzalez-Pinto, A., Querejeta, I. *et al.* (2004) Brief interventions for hazardous drinkers delivered in primary care are equally effective in men and women. *Addiction* **99**, 103–108.
- Beich, A., Gannik, D. and Malterud, K. (2002) Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. *BMJ* 325, 870–872.
- Beich, A., Thorsen, T. and Rollnick, S. (2003) Screening in brief intervention trials targeting excessive drinkers in general practice: systematic review and meta-analysis. *BMJ* 327, 536–540.
- Bertholet, N., Daeppen, J. B., Wietlisbach, V. et al. (2005) Reduction of alcohol consumption by brief alcohol intervention in primary care: systematic review and meta-analysis. Archives of Internal Medicine 165, 986–995.
- Bien, T. H., Miller, W. R. and Tonigan, J. S. (1993) Brief interventions for alcohol problems: a review. *Addiction* **88**, 315–335.
- Campbell, M., Fitzpatrick, R., Haines, A. et al. (2000) Framework for design and evaluation of complex interventions to improve health. BMJ 321, 694–696.
- Cohen, B. B. and Vinson, D. C. (1995) Retrospective self-report of alcohol consumption: test-retest reliability by telephone. *Alcoholism: Clinical and Experimental Research* 19, 1156–1161.
- Dawson, D. A. (1998) Volume of ethanol consumption: effects of different approaches to measurement. *Journal of Studies on Alcohol* **59**, 191–197.
- Edwards, A. G. and Rollnick, S. (1997) Outcome studies of brief alcohol intervention in general practice: the problem of lost subjects. *Addiction* **92**, 1699–1704.
- Flay, B. R. (1986) Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine* 15, 451–474.
- Fleming, M. F., Barry, K. L., Manwell, L. B. et al. (1997) Brief physician advice for problem alcohol drinkers. A randomized controlled trial in community-based primary care practices. *JAMA* 277, 1039–1045.
- Godwin, M., Ruhland, L., Casson, I. *et al.* (2003) Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Medical Research Methodology* **3**, 28.
- Gomberg, E. S. (1998) Alcoholic women in treatment: the question of stigma and age. *Alcohol and Alcoholism* **23**, 507–514.
- Gronbaek, M., Deis, A., Sorensen, T. I. *et al.* (1995) Mortality associated with moderate intakes of wine, beer, or spirits. *BMJ* **310**, 1165–1169.
- Hansen, L. J., Olivarius, N., Beich, A. et al. (1999) Encouraging GPs to undertake screening and a brief intervention in order to reduce problem drinking: a randomized controlled trial. Family Practice 16, 551–557.
- Heather, N. (2002) Effectiveness of brief interventions proved beyond reasonable doubt. *Addiction* **97**, 293–294.
- Heather, N. and Wallace, P. (2003) UK Research into brief interventions for excessive drinkers. Alcohol Concern Research Forum Papers (September issue). Alcohol Concern: London.
- Heather, N. (2006) WHO Collaborative Project on Identification and Management of Alcohol-Related Problems in Primary Health Care. Development of Country-wide Strategies for Implementing Early Identification and Brief Intervention in Primary Health Care, Report to the World Health Organization on Phase IV. World Health Organization, Department of Mental Health and Substance Abuse: Geneva: Available at: http://www.who.int/substance_abuse/publications/identification_management_alcoholproblems_phaseiv.pdf (accessed March 1, 2007).
- Kahan, M., Wilson, L. and Becker, L. (1995) Effectiveness of physician-based interventions with problem drinkers: a review. *Canadian Medical Association Journal* 152, 851–859.
- Kendrick, J. S., Zahniser, S. C., Miller, N. et al. (1995) Integrating smoking cessation into routine public prenatal care: the Smoking

- Cessation in Pregnancy project. *American Journal of Public Health* **85**, 217–222.
- Lemmens, P., Tan, E. S. and Knibbe, R. A. (1992) Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. *Journal of Studies on Alcohol* **53**, 476–486.
- Macpherson, H. (2004) Pragmatic clinical trials. *Complementary Therapies in Medicine* **12**, 136–140.
- Miller, W. R., Benefield, R. G. and Tonigan, J. S. (1993) Enhancing motivation for change in problem drinking: a controlled comparison of two therapist styles. *Journal of Consulting and Clinical Psychology* **61**, 455–461.
- Miller, W. R. and Rollnick, S. (1991) *Motivational Interviewing: Preparing People for Change.* Guilford Press: New York.
- Miller, W. R. and Rollnick, S. (2002) *Motivational Interviewing: Preparing People for Change*, 2nd edn. Guilford Press: New York.
- Miller, W. R. and Sanchez, V. C. (1993) *Motivating Young Adults for Treatment and Lifestyle Change*, In: Howard, G. (ed.) Issues in alcohol use and misuse by young adults. University of Notre Dame Press: Notre Dame.
- Moyer, A., Finney, J. W., Swearingen, C. E. *et al.* (2002) Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction* **97**, 279–292.
- Nitschke, K., Bennedsen, B. E., Kassow, P. *et al.* (1995) Incidence of alcohol abuse among patients with emergency admissions to a psychiatric department–Evaluation based on clinical and biochemical markers (article in Danish). *Ugeskrift for Laeger* **157**, 4567–4572.
- Peveler, R., Kendrick, T., Buxton, M. *et al.* (2005) A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. *Health Technology Assessment* **9**, 1–148.
- Poikolainen, K. (1999) Effectiveness of brief interventions to reduce alcohol intake in primary health care populations: a meta-analysis. *Preventive Medicine* **28**, 503–509.
- Rehm, J. (1998) Measuring quantity, frequency, and volume of drinking. *Alcoholism: Clinical and Experimental Research* 22, 4S-14S.
- Roland, M. and Torgerson, D. J. (1998) What are pragmatic trials? *BMJ* **316**, 285.
- Room, R. (1998) Measuring drinking practices: how far we've come and how far we need to go. *Alcoholism: Clinical and Experimental Research* **22**, 70S–75S.
- Saitz, R. (2005) Clinical practice. Unhealthy alcohol use. New England Journal of Medicine 352, 596–607.
- Saunders, J. B. and Aasland, O. G. (1987) WHO Collaborative Study on Identification and Treatment of Persons with Harmful Alcohol Consumption. Report on Phase 1: Development of a Screening Instrument. WHO: Geneva.
- Senft, R. A., Polen, M. R., Freeborn, D. K. et al. (1997) Brief intervention in a primary care setting for hazardous drinkers. American Journal of Preventive Medicine 13, 464–470.
- Sussman, A. L., Williams, R. L., Leverence, R. et al. (2006) The art and complexity of primary care clinicians' preventive counselling decisions: Obesity as a case study. Annals of Family Medicine 4, 327–333.
- Tunis, S. R., Stryer, D. B. and Clancy, C. M. (2003) Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 290, 1624–1632.
- US Preventive Services Task Force. (2004) Screening for Alcohol Misuse. Available at: http://www.ahrq.gov/clinic/uspstf/usps drin.htm (Accessed January 15, 2006).
- Wallace, P., Cutler, S. and Haines, A. (1988) Randomised controlled trial of general practitioner intervention in patients with excessive alcohol consumption. *BMJ* 297, 663–668.
- Whitlock, E. P. (2003) Alcohol screening in primary care. *BMJ* 327, E263–E264.
- Wilk, A. I., Jensen, N. M. and Havighurst, T. C. (1997) Meta-analysis of randomized control trials addressing brief interventions in heavy alcohol drinkers. *Journal of General Internal Medicine* 12, 274–283.

- Wilson, J. M. G. and Jungner, G. (1968) Principles and Practice of Screening for Disease.
 World Health Organization: (1993) The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and
- Diagnostic Guidelines. WHO: Geneva.

 World Health Organization. (2001) Collaborative Project on Identification and Management of Alcohol-related Problems in Primary Health Care. Available at: http://www.who-alcohol-phaseiv.net/ (Accessed March 15, 2006).
- World Health Organization. (2006) Screening and Brief Intervention for Alcohol Problems in Primary Health Care. Available at: http://www.who.int/substance_abuse/activities/sbi/en (Accessed January 15, 2006).
- Yarnall, K. S., Pollak, K. I., Ostbye, T. et al. (2003) Primary care: is there enough time for prevention? American Journal of Public *Health* **93**, 635–641.