We could not confirm some variables that had been predictors of postcesarean wound infection, namely duration of rupture of membranes,[1,9,14,23,24] and surgical duration.[10,20,26] Confounding effects were controlled sufficiently by multiple logistic regression in only two studies.[20,23] Prolonged rupture of membranes increased the likelihood of an infection ascending from vagina into uterine cavity. However, chorioamnionitis was closely related to prolonged rupture of membranes ($P < .001$). In a multivariate model this strong predictor would have masked the hypothesized association between prolonged rupture of membranes and postcesarean infection.

Although the cesareans that lasted longer than 1 hour had 2.4 times the risk of postoperative infection, by univariate analysis, a larger sample is needed to confirm its independent predictive role. Assuming an α of .05, power of .80, and assumed infection rate in controls of 8.5%, at least 264 subjects would be needed for each study arm.

References

1. Ott WJ. Primary cesarean section: Factors related to postpartum infection. Obstet Gynecol 1981;57:171–6.
2. Garner JS, Jarvis WR, Emori TG, Horan TC, Hughes JM. CDC definitions for nosocomial infections. 1988. Am J Infect Control 1988;16:128–40.
3. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections. 1992: A modification of CDC definitions of surgical wound infections. Infect Control Hosp Epidemiol 1992;13:606–8.
4. Wong ES. Surgical site infection. In: Mayhall CG, ed. Hospital epidemiology and infection control. Baltimore: Williams & Wilkins, 1996:154–75.
5. Yalcin N, Bakir M, Doknetas I, Sabir N. Postoperative wound infections. J Hosp Infect 1995;29:305–9.
6. Starling CE, Cutto BR, Pinheiro SM. Applying the Centers for Disease Control and Prevention and National Nosocomial Surveillance system methods in Brazilian hospitals. Am J Infect Control 1997;25:303–11.
7. Eltahawy AT, Mokhtar AA, Khalar RMF, Bahnassy AA. Postoperative wound infection at a university hospital in Jeddah, Saudi Arabia. J Hosp Infect 1992;21:79–83.
8. Owens WD, Felts JA, Spitznagel EL. ASA physical status classifications: A study of consistency of ratings. Anesthesiology 1978;49: 239–43.
9. Emmons SL, Krohn M, Jackson M, Eschenbach DA. Development of wound infections among women undergoing cesarean section. Obstet Gynecol 1988;72:559–64.
10. Hagglund L, Christensen K, Christensen P, Kamme C. Risk factors in cesarean section infection. Obstet Gynecol 1983;62:145–50.
11. Cunningham FG, MacDonald PC, Gant NF, Levenio KJ, Gilstrap LC III. Puerperal infection. In: Williams obstetrics. 19th edition. Norwalk, Connecticut: Appleton & Lange. 1993:627–42.
12. Nielsen TF, Hokegard KH. Cesarean section and intraoperative surgical complications. Acta Obstet Gynecol Scand 1984;63:103–8.
13. Baskett TF, McMillen RM. Cesarean section: Trends and morbidity. Can Med Assoc J 1981;125:723–6.
14. Nielsen TF, Hokegard KH. Postoperative cesarean section morbidity: A prospective study. Am J Obstet Gynecol 1983;146:911–5.
15. Eisenkop SM, Richman R, Platt LD, Paul RH. Urinary tract injury during cesarean section. Obstet Gynecol 1982;60:591–6.
16. Guldholt I, Espersen T. Maternal febrile morbidity after cesarean section. Acta Obstet Gynecol Scand 1987;66:675–9.
17. Gravel-Tropper D, Oxley C, Memish Z, Garber GE. Underestimation of surgical site infection rates in obstetrics and gynecology. Am J Infect Control 1995;23:22–6.
18. Garibaldi RA, Cushing D, Lerer T. Predictors of intraoperative-acquired surgical wound infections. J Hosp Infect 1991;18(Suppl A):289–98.
19. Newton ER, Prihoda TJ, Gibbs RS. A clinical and microbiologic analysis of risk factors for puerperal endometritis. Obstet Gynecol 1990;75:402–6.
20. Suonio S, Saarikoski S, Vohlonen I, Kauhanen O. Risk factors for fever, endometritis and wound infection after abdominal delivery. Int J Gynaecol Obstet 1989;29:135–42.
21. Dinsmoor MJ, Gibbs RS. Previous intra-amniotic infection as a risk factor for subsequent peripartal uterine infections. Obstet Gynecol 1989;74:299–301.
22. Hill JA, Devoe LD, Bryans CI Jr. Frequency of asymptomatic bacteriuria in preeclampsia. Obstet Gynecol 1986;67:529–32.
23. Pelle H, Jepsen OB, Larsen SO, Bo J, Christensen F, Dreisler A, et al. Wound infection after cesarean section. J Infect Control 1986;7: 456–61.
24. Marteris M, Kolrud B, Faro S, Maccato M, Hammill H. Development of wound infection or separation after cesarean delivery. Prospective evaluation of 2,431 cases. J Reprod Med 1995;40:171–5.
25. Sheikh GN. Observation of maternal weight behavior during the puerperium. Am J Obstet Gynecol 1971;111:244–50.
26. Chang PL, Newton ER. Predictors of antibiotic prophylactic failure in post-cesarean endometritis. Obstet Gynecol 1992;80:117–22.

Address reprint requests to:
Thach Son Tran, MD, PhD
Postoperative Department
Hungvuong Obstetric and Gynecological Hospital
128 Hungvuong Street, District 05
Ho Chi Minh City
Vietnam
E-mail: transon_thach@hotmail.com

Received May 28, 1999.
Received in revised form August 6, 1999.
Accepted August 19, 1999.

Biotypes of oral *Candida albicans* isolated from AIDS patients and HIV-free subjects in Thailand

R. Teanpaisan[1]
W. Nittayananta[2]
V. Chongsuvivatwong[2]

[1]Department of Stomatology, Faculty of Dentistry and
[2]Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand

Abstract: This study was conducted to examine biotypes and antifungal susceptibility patterns of oral *Candida albicans* isolated from HIV-infected patients, HIV-free patients with candidiasis and healthy subjects. All isolates were biotyped using a typing system based on enzyme profiles, carbohydrate assimilation patterns and boric acid resistance. Thirty-eight biotypes were found amongst 218 oral *C. albicans* isolates. The major biotype found was A1S, which accounted for 32.6% of all isolates, and this biotype was the most common in all groups. There was a greater variety of biotypes of *C. albicans* in the HIV-infected group than in the other groups; however, there was no statistically significant difference between the groups. The minimum inhibitory concentrations (MICs) of a total of 118 isolates were determined for amphotericin B and for ketoconazole using the National Committee for Clinical Laboratory Standards (NCCLS) broth macrodilution method and the E-test. When the antifungal susceptibility patterns among the groups were compared, a statistically significant difference was found only with amphotericin B. The median MIC of amphotericin B in the HIV-infected group was higher than in the healthy group ($P=0.013$, NCCLS method; $P=0.002$, E-test). However, this difference in sensitivity was not restricted to any sub-type investigated. Our results showed that the biotype patterns of *C. albicans* isolates that colonize HIV-infected patients are similar to those of HIV-free subjects, and there is no relationship between antifungal susceptibility patterns and the biotypes.

Key words: biotypes; *Candida albicans*; HIV infection; Thailand

J Oral Pathol Med 2000; **29**: 193–9

A number of reports have revealed that the frequency of isolation of *Candida* and clinical signs of oral candidiasis increase with advancing HIV infection (1–4). Although *C. albicans* is the most common of the *Candida* species isolated in these studies, relatively few details of the pathogenic features of this organism and AIDS-associated oral candidiasis are known. In healthy humans, *C. albicans* is present as part of the normal flora of the oral cavity and gastrointestinal tract; however, in immunosuppressed patients it can cause severe mucosal or invasive disease (5). The high incidence of mu-

Correspondence to:
R. Teanpaisan
Department of Stomatology, Faculty of Dentistry, Prince of Songkla University, Hat Yai, P.O. Box 17 Khor Hong, Songkhla 90112, Thailand

Accepted for publication October 13, 1999

oral candidiasis in patients with AIDS may be due to infection with the same strains that are non-pathogenic in healthy subjects but that become pathogenic in AIDS patients due to impaired host defence mechanisms. Alternatively, it may be due to infection with unique or more virulent strains. In addition, the significance of oral candidiasis as a disease entity in HIV-related immunosuppression is its frequent recurrence (6, 7). The mechanism behind the ability of this fungus to cause recurrent disease is unknown. It is postulated that one factor that may influence this is a decrease in susceptibility to antifungal agents (7, 8), and this may be related to certain types of *C. albicans* (9). Because of this, many attempts have been made to search for particularly virulent types of *C. albicans* using a number of techniques, such as serotyping (10), biotyping (11), morphotyping (12, 13) and genotyping (14, 15). On account of the difficulties with some of these techniques, a standardised, simple and highly specific biotyping system developed by Williamson et al. (11), which is technically undemanding and utilises relatively inexpensive commercially available quality controlled media, has been increasingly used in recent years (16–19).

In a previous paper (4) we showed that there is an association between candidal load and HIV infection. The aims of the present study were to determine whether any sub-strains of *C. albicans* isolated from HIV patients, HIV-free subjects with candidiasis and healthy subjects are particularly associated with health or disease using the biotyping method of Williamson et al. (11), and to compare isolates for their susceptibility to two commonly used antifungal agents (amphotericin B and ketoconazole).

Material and methods

Sources of *C. albicans*

A total of 218 isolates of *C. albicans* were included in this study, which comprised 82 isolates from salivary samples of 15 HIV-infected patients, 76 isolates from 15 HIV-free patients with oral candidiasis and 60 isolates from 16 healthy subjects. The specimen collection and methods used for fungal cultivation and identification have been described previously (4). Briefly, an oral rinse specimen was obtained from each subject using 10 ml sterile phosphate buffered saline according to the method of Samaranayake et al. (20). Colonies showing yeast-like morphology within the 48–72 h incubation period were selected for study. All colonies showing variation in morphology, as well as identical colonies, were selected from the same isolation plate. The number of colonies chosen depended on the density of growth recovered on the primary culture plate. For example, usually 10–20 colonies from 20–40 colonies on the same plate were selected for identification and biotyping, but when recovery was lower (e.g., from healthy subjects) fewer colonies could be sampled. All isolates were identified by using production of chlamydospores, production of germ tubes and carbohydrate assimilation with API 20 C AUX (Bio Merieux, France).

Biotyping of *C. albicans* isolates

All isolates of *C. albicans* were biotyped using the method of Williamson et al. (11), which employs two commercially available kits, [API ZYM and API 20 C AUX (Bio Merieux)], and a boric acid resistance test. In brief, the API ZYM system evaluates the enzyme activity of the isolates by means of a set of 19 enzyme substrates contained in a tray of miniaturised plastic cupules. After inoculation of a standard suspension of the organism and incubation for 4 h at 37°C, the colour reactions in each cupule were read according to the manufacturer's instructions. The API 20 C AUX system utilises the ability of *C. albicans* isolates to assimilate 19 different carbohydrates as sole sources of carbon. The results were determined by comparison of the opacity in the test and control cupules. Finally, the boric acid resistance test assesses the sensitivity of the isolates to 1.8 mg/ml of boric acid incorporated into an agar medium.

Antifungal susceptibility testing

A total of 118 isolates of *C. albicans* were chosen to represent strains with either different or the same colony morphology and biotype. These included 52 isolates from the HIV-infected group, 33 isolates from the HIV-free candidiasis group and 33 isolates from the healthy group. For each isolate, minimal inhibitory concentrations (MICs) for amphotericin B and ketoconazole were determined using the NCCLS macrodilution method (21) and the E-test strip (AB Biodisk, Sweden).

The NCCLS broth dilution method was performed according to NCCLS document M27-P (21). A working suspension of the inoculum was made by a 1:100 dilution of the 0.5 McFarland standard yeast suspension in 0.85% saline followed by a 1:20 dilution in RPMI broth. Two-fold dilutions of the antifungal agents from 64 to 0.015 μg/ml were prepared and inoculated with the working suspension. The tubes were incubated at 37°C for 48 h. The MIC was read as the concentration that inhibited growth (amphotericin B) or produced an 80% reduction of turbidity in comparison with a drug-free control (ketoconazole). For the E-test, colonies of each yeast were suspended in saline to produce a turbidity equivalent to a 0.5 McFarland standard. The inoculum was swabbed on to Sabouraud's dextrose agar and allowed to dry for 10–15 min before each of the antifungal E-test strips was applied. Plates were read at 24 h and

Table 1. Biotype profiles of oral *Candida albicans* isolated from HIV infected patients, HIV-free patients with candidiasis and healthy subjects

Biotypes	HIV-infected group No. (%)	HIV-free with candidiasis No. (%)	Healthy subjects No. (%)	Total
A1R	3 (3.6)	0	1 (1.7)	4 (1.8)
A1S	26 (31.7)	25 (32.8)	20 (33.3)	71 (32.6)
A4R	1 (1.2)	0	1 (1.7)	2 (0.9)
A4S	6 (7.3)	6 (7.8)	5 (8.3)	17 (7.8)
A65*	0	0	1 (1.7)	1 (0.5)
A7S*	1 (1.2)	0	0	1 (0.5)
A8S	4 (4.9)	2 (2.6)	0	6 (2.6)
A14R	1 (1.2)	0	0	1 (0.5)
A17S	0	1 (1.3)	1 (1.7)	2 (0.9)
A18S	0	2 (2.6)	3 (5.0)	5 (2.3)
A19S*	1 (1.2)	0	0	1 (0.5)
A20S*	1 (1.2)	0	0	1 (0.5)
A23S*	0	1 (1.3)	1 (1.7)	2 (0.9)
A24S*	0	0	1 (1.7)	1 (0.5)
B1R	0	6 (7.9)	5 (8.3)	11 (5.0)
B1S	4 (4.9)	6 (7.9)	7 (11.7)	17 (7.8)
B2S*	0	0	1 (1.7)	1 (0.5)
B4R*	1 (1.2)	2 (2.6)	0	3 (1.4)
B4S*	13 (15.9)	12 (15.8)	7 (11.7)	32 (14.7)
B6S*	2 (2.4)	1 (1.3)	0	3 (1.4)
B15S*	0	0	1 (1.7)	1 (0.5)
B16S*	0	0	1 (1.7)	1 (0.5)
B16R*	0	2 (2.6)	0	2 (0.9)
B15S*	1 (1.2)	0	0	1 (0.5)
B19S*	1 (1.2)	0	0	1 (0.5)
B20S*	2 (2.4)	0	0	2 (0.9)
B21S*	1 (1.2)	1 (1.3)	0	2 (0.9)
B22S*	1 (1.2)	0	0	1 (0.5)
B24S*	1 (1.2)	0	0	1 (0.5)
C1S	1 (1.2)	2 (2.6)	0	3 (1.4)
D1R	1 (1.2)	1 (1.3)	0	2 (0.9)
D1S	6 (7.3)	2 (2.6)	0	8 (3.7)
D6S*	0	2 (2.6)	1 (1.7)	3 (1.4)
D8S	3 (3.6)	0	0	3 (1.6)
D24S*	0	0	1 (1.7)	1 (0.5)
F1R	0	0	1 (1.7)	1 (0.5)
F4S*	0	0	1 (1.7)	1 (0.5)
H4S*	0	2 (2.6)	0	2 (0.9)
Total	82 (100)	76 (100)	60 (100)	218 (100)
*New biotypes	26 (31.7)	23 (30.3)	16 (26.7)	65 (29.8)

48 h according to the E-test technical guide for antifungal susceptibility testing.

Results

The biotyping system employed in this study utilized three tests: the API ZYM test (the first letter of the code), the API 20 C system (the middle digit of the code), and resistance or sensitivity to boric acid, denoted as either R or S (the last letter of the code). The reproducibility of the results obtained by these methods has not been tested on a broad scale, but some isolates chosen at random were tested on multiple occasions and have shown clearly reproducible results. A total of 38 biotypes was found among the 218 oral *C. albicans* isolates (Table 1). The major biotype, A1S, accounted for 32.6% of the isolates, and this biotype was commonly found in HIV-infected patients, HIV-free candidiasis subjects and healthy subjects (31.7%, 32.8%, and 33.3%, respectively). The second most common biotype was B4S, which represented 14.7% of total isolates. The other biotypes found were A4S, B1S and B1R (7.8%, 7.8% and 5.0%, respectively). When the number of different biotypes in each of the HIV-infected patients, HIV-free with candidiasis and healthy subjects was compared, the first group had a higher number of biotypes than the others (Fig. 1); however, this difference was not statistically significant (Kruskal-Wallis test).

A total of 118 strains of *C. albicans* obtained from HIV-infected patients, HIV-free candidiasis subjects and healthy subjects were compared for their MICs against amphotericin B and ketoconazole using the NCCLS macrodilution method and the E-test, and the results are shown in Table 2. Generally, results obtained by the NCCLS broth method were largely in agreement with those obtained by the E-test. Pearson correlation coefficients of both methods for amphotericin B and ketoconazole were 0.89 and 0.57, respectively. The median MIC for amphotericin B in isolates from the HIV-infected group was statistically significantly higher than that of isolates from the healthy subjects ($P=0.013$, NCCLS method; $P=0.002$, E-test). Also, there were statistically significant differences in the mean MIC for amphotericin B between the HIV-infected group and the HIV-free subjects with candidiasis ($P=0.01$, NCCLS method) and between the HIV-free candidiasis group and the healthy group ($P<0.03$, E-test). When the patients' history of antifungal therapy was taken into account, it showed that there was no significant difference between the MIC and whether or not a patient had taken amphotericin B previously. The mean MICs of ketoconazole among

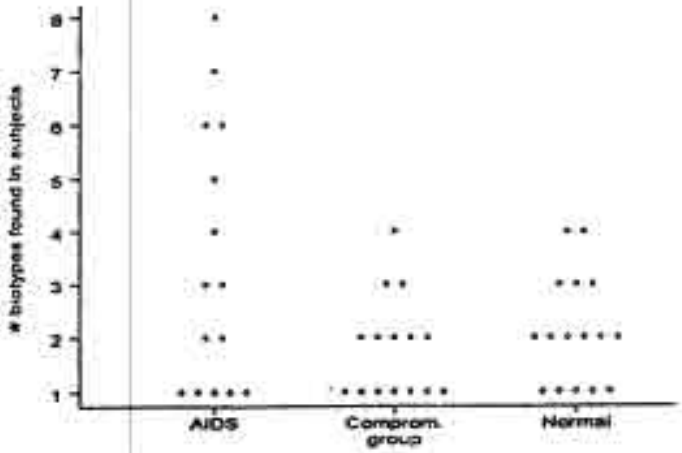


Fig. 1. Number of biotypes identified in the AIDS group, HIV-free patients with candidiasis (compom.) group and normal subjects.

Table 2. Median of MICs (μg/ml) of amphotericin B and ketoconazole

Group	Amphotericin B				Ketoconazole	
	NCCLS	*P*-value	E-test	*P*-value	NCCLS	E-test
HIV-infection:	0.500 (0.062–4.0)	0.013*	0.315 (0.032–>32)	0.002*	0.062 (0.015–16.0)	0.094 (0.012–>32)
Taking antifungals	0.500 (0.250–1.0)		0.380 (0.032–>32)		0.062 (0.031–1.0)	0.094 (0.012–>32)
No antifungals	0.500 (0.125–0.500)		0.250 (0.190–0.500)		0.062 (0.015–4.0)	0.125 (0.032–0.500)
HIV-free candidiasis:	0.250 (0.562–1.0)	0.01[a]	0.250 (0.038–0.500)	0.49*	0.062 (0.015–8.0)	0.094 (0.032–>32)
Taking antifungals	0.310 (0.125–0.500)	"	0.380 (0.125–0.500)		0.062 (0.062–0.125)	0.079 (0.032–0.094)
No antifungals	0.250 (0.062–1.0)		0.250 (0.038–0.500)		0.062 (0.015–8.0)	0.125 (0.032–>32)
Healthy subjects	0.250 (0.031–1.0)	0.96*	0.190 (0.032–0.500)	0.03*	0.062 (0.015–8.0)	0.064 (0.047–>32)

* Median MICs of HIV group vs healthy subjects.
[a] Median MICs of HIV group vs HIV-free candidiasis group.
* Median MICs of HIV-free candidiasis patients vs healthy subjects.

isolates from the three groups showed no significant differences using either the NCCLS method or the E-test.

Discussion

Among the numerous AIDS-associated oral diseases, oral candidiasis is the most frequent, with up to 90% of HIV-infected patients being affected (22, 23). The clinical syndrome is not life-threatening but it is painful and its recurrent nature makes it of importance. Oral candidiasis is caused mainly by *C. albicans* but its pathogenesis is still unclear. The condition of the patient is probably the major factor governing the development of clinical candidiasis and this is often associated with immunodeficiencies (24). However, as recently shown for many other microbial pathogens, the possibility that certain strains or groups of strains are more likely to be involved in clinical disorders cannot be excluded (25). Whether or not HIV patients are colonized with selected strains of *C. albicans* has been a matter of debate. Using DNA fingerprinting, the results of two studies showed that no particular strain was associated with HIV-infected patients and that *C. albicans* populations from the oral cavities of HIV-infected and HIV-negative people have a similarly disparate clonal origin (26, 27). Some researchers have used a DNA probe to track the *C. albicans* isolates from oral lesions in HIV-seropositive individuals, and their data suggest that each patient carries a unique strain of *C. albicans*. Furthermore, it was shown that the strains present during both symptomatic and asymptomatic states of candidiasis were the same (28, 29). However, others have produced evidence that there is increased genetic variation of *C. albicans* isolates in HIV-infection compared to controls (30, 31). These differences were noted when the appropriate molecular technique coupled with appropriate analyses were used (31). Also, Sweet et al. (32, 33) showed that more biotypes of *C. albicans* were present in HIV/AIDS groups than in control subjects, and that almost all *Candida* species isolated from HIV subjects adhered to buccal epithelial cells in higher numbers than did those strains isolated from controls.

Here, we have used the biotyping method of Williamson et al. (11) to evaluate the biotypes of isolates from a range of patient groups. It is known that genetic typing methods provide more sensitive and specific means to discriminate among isolates. However, biotyping has the advantages of being simple to perform, technically undemanding and inexpensive compared to molecular techniques. In addition, it also has the advantage of allowing more meaningful comparison of the present studies with those of previous studies of a similar nature. The first study of *C. albicans* biotypes in HIV-infected patients was conducted by Korting et al. (34) and employed the API 20 C (carbohydrate assimilation) system. Their results showed that from a total of 61 oral *C. albicans* strains isolated from HIV-infected individuals, with or without signs of candidiasis, the majority (64%) of isolates belonged to group 1. However, there was no detail given of the incidence and proportions of each of the biotypes between the groups with and without signs of clinical candidiasis, and no HIV-negative subjects were included in the study. Since the API 20 C carbohydrate assimilation system has a relatively poor discriminatory power, API ZYM and boric acid sensitivity tests were added to complement the API 20 C profiles (11). Results from previous studies of others (16–19) have demonstrated that this system serves to further differentiate the biotypes into smaller sub-groups. Using this biotyping system, Tsang et al. (17) showed that there are many different sub-strains of oral *C. albicans* in HIV-infected patients. However, no data on healthy subjects were provided. Our present results showed that there were 38 biotypes among 216 strains. When the results were compared among the groups, HIV-infected patients had more sub-types (1–8) than HIV-free candidiasis patients (1–4) or healthy subjects (1–4), and there were no

significant differences in the biotypes between the three groups. The results of statistical analysis suggested that this is maybe affected by the small sample size in this study. Thus, larger numbers of patients will need to be studied for further confirmation.

With regard to the geographic distribution of the different biotypes, it has been shown that some biotypes are globally prevalent. However, almost one-third of the biotypes reported here have not been previously described and may reflect geographical exclusivity. Tsang et al. (17) has reported that A1R (18%) and A1S (11%) are the most common biotypes among the oral *C. albicans* isolates derived from HIV-infected patients in Hong-Kong, Australia, England and Germany. Another previous study in healthy individuals in Britain found that A1R and A1S were also the commonest biotypes, accounting for 23% and 26% of the total isolates, respectively (11). The biotypes A1S and J1S have been found to predominate in China (18) and in Tanzania (19). Our results concur with the findings in China and Tanzania that the most common biotype is A1S, accounting for 32.6% in all groups investigated. It is noted that in previous studies (17–19), only a single representative isolate was selected from each culture plate, whereas multiple colonies from each plate were collected in the present study. It is assumed that the colonies examined represent the predominant strains present on the plate. No clear explanation for the widespread over-representation of such biotypes has been given. It may be hypothesized that they are better adapted than other sub-types to life on or in the human body. They may also be more easily transmitted between humans than are other sub-types. When compared with previous data of new biotypes from the foregoing countries, our results showed that almost one-third (65 of 218) of isolates were previously undescribed new biotypes. As the previous reports were from Scotland (11), Germany, Australia, England and Hong-Kong (17), China (18) and Tanzania (19), it is likely that there are geographical variations in *C. albicans* biotypes.

Up to now, there has been no general agreement about a standardized method of *in vitro* antifungal susceptibility testing, since results have shown great inter- and intra-laboratory variations. The NCCLS has established a broth macrodilution method as a reference method for antifungal susceptibility testing; however, it is labour-intensive and time-consuming. We are in agreement with the previous studies that the E-test appears to be equivalent to the NCCLS reference macrobroth method for testing susceptibility of *Candida* species to azole antifungal agents (35, 36). Wanger et al. (37) has concluded that the E-test is comparable to the NCCLS method for testing of susceptibility to amphotericin B and fluconazole; in addition, the E-test appears to be superior for the detection of resistance to amphotericin B. Our results have shown that the NCCLS reference macrodilution method and the E-test gave very similar results for amphotericin B and showed moderate agreement for ketoconazole. Generally, the endpoints obtained were identical or different by no more than two two-fold dilutions. However, when the MIC level is high, the results of both methods show a greater difference; if the MICs given by the NCCLS method are more than 1–4 μg/ml, they would be >32 μg/ml by the E-test. This may be due to the problem of diffusion of the agent through the agar medium necessary for the E-test. In the present study, we have shown that susceptibility of our isolates to amphotericin B was significantly different between the patient groups but that there was no difference in sensitivity to ketoconazole. It was found that isolates from the AIDS group were more resistant to amphotericin B than were isolates from the HIV-free candidiasis group and the healthy group. This resistance is not associated with a history of amphotericin B therapy or restricted to any sub-type investigated (data not shown). Gallagher et al. (38) have shown that phenotypically switched variants of *C. albicans* can develop decreased azole susceptibility even though these strains remained genetically identical. The results of McCullough et al. (39) showed that the same *C. albicans* genotypes tended to persist during the course of disease progression, but that the colonial morphologies of the isolates changed. Also, Soll et al. (40) found that despite a high frequency of phenotypic switching by *C. albicans*, nucleic acid hybridization of DNA from multiple phenotypes from a single culture site consistently yielded identical genotypes. These observations have shown that it is not necessary to have alterations in the type of strain present for there to be changes in the drug susceptibility of *C. albicans*. This may explain our finding of increased resistance to amphotericin B among *C. albicans* strains isolated from HIV-infected patients but no difference in biotypes. However, further work on genetic analysis is required to clarify this.

To conclude, it is worth emphasizing here that the present data are the first base-line information for studies of oral candidal infection in different cohorts (AIDS patients, HIV-free healthy subjects and HIV-free patients with candidiasis) of the Thai population. Our results show that the biotype patterns of *C. albicans* that colonize AIDS patients are similar to those in normal Thai subjects. This suggests that there may be no particular biotypes of *C. albicans* linked with specific clinical characteristics. Similarly, the oral *C. albicans* isolates showing *in vitro* susceptibility to amphotericin B and ketoconazole were not restricted to any sub-group investigated. However, the concept of phenotypic switching of *C. albicans* among HIV-infected patients could not be excluded, and the resistance of these isolates to amphotericin B needs further explanation.

References

1. Torssander J, Mörfeldt-Manson L, Biberfeld G, Karlsson A, Putkonen PD, Wasserman J. Oral *Candida albicans* in HIV infection. *Scand J Infect Dis* 1987; 19: 291–5.
2. Schmidt-Westhausen A, Schiller RA, Pohle HD, Reichart PA. Oral *Candida* and *Enterobacteriaceae* in HIV-1 infection: correlation with clinical candidiasis and antimycotic therapy. *J Oral Pathol Med* 1991; 20: 469–72.
3. Samaranayake LP. Oral mycoses in HIV infection. *Oral Surg Oral Med Oral Pathol* 1992; 73: 171–80.
4. Teanpaisan R, Nittayananta W. Prevalence of *Candida* species in AIDS patients and HIV-free subjects in Thailand. *J Oral Pathol Med* 1998; 27: 4–7.
5. Odds FC. *Candida* and candidiasis. A review and bibliography. 2nd edn. Philadelphia: W. B. Saunders, 1988.
6. Glatt AE, Chirgwin K, Landesmann SH. Treatment of infections associated with human immunodeficiency virus. *N Engl J Med* 1988; 318: 1429–48.
7. Samaranayake LP, Holmstrup P. Oral candidiasis and human immunodeficiency virus infection. *J Oral Pathol Med* 1989; 18: 554–64.
8. Larsen RA. Azoles and AIDS. *J Infect Dis* 1990; 162: 727–30.
9. Odds FC, Abbott AB, Stiller RL, Scholer HJ, Polak A, Stevens DA. Analysis of *Candida albicans* phenotypes from different geographical and anatomical sources. *J Clin Microbiol* 1983; 18: 849–57.
10. Hazenclever HF, Mitchell WO. Antigenic studies of *Candida*. Observation of two antigenic groups in *Candida albicans*. *J Bacteriol* 1971; 82: 570–3.
11. Williamson MI, Samaranayake LP, MacFarlane TW. A new simple method for biotyping *Candida albicans*. *Microbios* 1987; 51: 159–67.
12. Phongpaichit S, MacKenzie DWR, Fraser C. Strain differentiation of *Candida albicans* by morphotyping. *Epidemiol Infect* 1987; 99: 421–8.
13. Oliver AJ, Reade PC. Morphotypes of oral isolates of *Candida albicans* from patients infected with the human immunodeficiency virus. *J Med Vet Mycol* 1993; 31: 289–97.
14. Magee BB, D'Souza TM, Magee PT. Strain and species identification by restriction fragment length polymorphisms in the ribosomal DNA repeat of *Candida* species. *J Bacteriol* 1987; 169: 1639–43.
15. Matthews R, Burnie J. Assessment of DNA fingerprinting for rapid identification of outbreaks of systemic candidiasis. *Br Med J* 1989; 298: 334–7.
16. Rams TE, Slots J. *Candida* biotypes in human adult periodontitis. *Oral Microbiol Immunol* 1991; 6: 191–2.
17. Tsang PCR, Samaranayake LP, Philipsen HP, et al. Biotypes of oral *Candida albicans* isolates in human immunodeficiency virus-infected patients from diverse geographic locations. *J Oral Pathol Med* 1995; 24: 32–6.
18. Xu YY, Samaranayake LP. Oral *Candida albicans* biotypes in Chinese patients with and without oral candidosis. *Archs Oral Biol* 1995; 40: 577–9.
19. Matee MI, Samaranayake LP, Scheutz F, Simon E, Lyamuya EF, Mwinula J. Biotypes of oral *Candida albicans* isolates in a Tanzanian child population. *APMIS* 1996; 104: 623–8.
20. Samaranayake LP, MacFarlane TW, Lamey PJ, Ferguson MM. A comparison of oral rinse and imprint sampling techniques for the detection of yeast, coliform and *Staphylococcus aureus* carriage in the oral cavity. *J Oral Pathol* 1986; 15: 386–8.
21. National Committee for Clinical Laboratory Standards. Reference method for broth dilution antifungal susceptibility testing of yeasts. Proposed standard M 27-P. Villanova, PA: National Committee for Clinical Laboratory Standards, 1992.
22. Phelan JA, Saltzman BR, Friedland GH, Klein RS. Oral findings in patients with acquired immunodeficiency syndrome. *Oral Surg Oral Med Oral Pathol* 1987; 64: 50–6.
23. Coleman D, Russell R, Harwood M, Mulcahy F, Shanley D. Clinical and microbiological analysis of oral candidiasis in HIV positive patients. *J Dent Res* 1989; 68: 893.
24. Wade JC. Epidemiology of *Candida* infections. In: Hodey GP, ed. *Candidiasis: pathogenesis, diagnosis and treatment*. New York: Raven Press, Ltd., 1993; 65–107.
25. Soll DR, Galask R, Schmid J, Hanna C, Mac K, Morrow B. Genetic dissimilarity of commensal strains of *Candida* spp. carried in different anatomical locations of the same healthy women. *J Clin Microbiol* 1991; 29: 1702–10.
26. Whelan WL, Kirsch DR, Kwon-Chung KJ, Wahl SM, Smith PD. *Candida albicans* in patients with the acquired immunodeficiency syndrome: absence of a novel or hypervirulent strain. *J Infect Dis* 1990; 162: 513–8.
27. Powderly WG, Robinson K, Keath EJ. Molecular typing of *Candida albicans* isolated from oral lesions of HIV-infected individuals. *AIDS* 1992; 6: 81–4.
28. Miyasaki SH, Hicks JB, Greenspan D, et al. The identification and tracking of *Candida albicans* isolates from oral lesions in HIV-seropositive individuals. *J Acquir Immune Defic Syndr Hum Retrovirol* 1992; 5: 1029–36.
29. Schmid J, Odds FC, Wiselka MJ, Nicholson KG, Soll DR. Genetic similarity and maintenance of *Candida albicans* strains from a group of AIDS patients, demonstrated by DNA fingerprinting. *J Clin Microbiol* 1992; 30: 935–41.
30. Anthony RM, Midgley J, Sweet SP, Howell SA. Multiple strains of *Candida albicans* in the oral cavity of HIV positive and HIV negative patients. *Microbial Ecol Health Dis* 1995; 8: 23–30.
31. Challacombe SJ, Muir J, Howell SA, Sweet SP. Genetic variability of *Candida albicans* in HIV infection. *Microbial Ecol Health Dis* 1995; 8: 63–70.
32. Sweet SP. Oral candidiasis in HIV infection. *J Dent Res* 1994; 73: 792 (abstr).
33. Sweet SP, Cookson S, Challacombe SJ. *Candida albicans* isolates from HIV-infected and AIDS patients exhibit enhanced adherence to epithelial cells. *J Med Microbiol* 1995; 43: 452–7.
34. Korting HC, Ollert M, Georgii A, Froschl M. *In vitro* susceptibilities and biotypes of *Candida albicans* isolates from the oral cavities of patients infected with human immunodeficiency virus. *J Clin Microbiol* 1988; 26: 2626–31.
35. Espinel-Ingroff A. E-test for antifungal susceptibility testing of yeasts. *Diagn Microbiol Infect Dis* 1994; 19: 217–20.
36. Colombo AL, Barchiesi F, McGough DA, Rinaldi MG. Comparison of E-test and National Committee for Clinical Laboratory Standards broth macrodilution method for azole antifungal susceptibility testing. *J Clin Microbiol* 1995; 33: 535–40.
37. Wanger A, Mills K, Nelson PW, Rex JH. Comparison of E-test and National Committee for Clinical Laboratory Standards broth macrodilution method for antifungal susceptibility testing: enhanced ability to detect amphotericin B-resistant *Candida* isolates. *Antimicrob Agents Chemother* 1995; 39: 2520–22.
38. Gallagher PJ, Bennett DE, Herman NC, et al. Reduced azole susceptibility of oral isolates of *Candida albicans* from HIV-positive patients and a

derivative exhibiting colony morphology variation. *J Gen Microbiol* 1992; 138: 1901–11.
McCullough M, Ross B, Reade PC. Oral *Candida albicans* from patients infected with the human immunodeficiency virus and characterization of a genetically distinct subgroup of *Candida albicans*. *Aus Dent J* 1995; 40: 91–7.
Soll DR. A molecular approach to the role of switching in oral candidiasis. In: Greenspan JS, Greenspan D, eds. *Oral manifestations of HIV infection*. Chicago: Quintessence Publishing Co Inc, 1995; 93–102.

Acknowledgements

This study was supported by a grant from Prince of Songkla University. The third author is supported by the Thailand Research Fund. We are very grateful to Dr. C. W. I. Douglas for his valuable advice.

COMMUNITY-BASED SELF-REPORTED SYMPTOMS OF ANTEPARTUM MORBIDITIES; THE HEALTH BURDEN AND CARE-SEEKING PATTERNS OF RURAL BANGLADESHI WOMEN

Mahbub-E-Elahi Khan Chowdhury[1], Halida Hanum Akhter[1] and Virasakdi Chongsuvivatwong[2]

[1]Bangladesh Institute of Research for Promotion of Essential and Reproductive Health and Technologies (BIRPERHT); [2]Epidemiology Unit, Faculty of Medicine, Prince of Songkla University (PSU), Hat Yai, Thailand

Abstract. In Bangladesh there is a dearth on information relating to complications during pregnancy. We followed up 1,019 pregnant women in rural Bangladesh sampled from all the 4 old administrative divisions of the country. Trained female interviewers visited households of the pregnant women at four-week intervals and interviewed them for their current pregnancy-related complications. Out of a total of 3,812 antepartum visits the percentage of reported symptoms of bleeding, fits and convulsions, excessive vomiting, fever >3 days, urinary problems, palpitations and symptomatic anemia were 0.3, 0.7, 1.4, 4.0, 26.8, 46.5 and 78.3 respectively. Morbidities were considered to cause a health burden if they imposed constraints in daily activities of the pregnant women and they were weighted according to intensity of the constraint. For each morbidity, the mean intensity of burden per episode and the population burden per 1,000 person months of observation of all the women were calculated. For common sustaining morbidities like symptomatic anemia and urinary problems the population burden was much heavier than that for more serious but rare morbidities like bleeding and convulsions. Among the visits in which the women had any symptoms, the percentages of care-seeking for less frequently reported morbidities such as fits and convulsions, bleeding, fever >3 days, excessive vomiting were about 74, 50, 34 and 33% respectively, whereas those for more commonly reported complications such as urinary problems, symptomatic anemia and palpitations were less than 20%. Care for these morbidities was mostly sought from untrained providers.

INTRODUCTION

More than half a million women die every year due to pregnancy complications or childbirth and about 99% of these deaths occur in the developing countries (WHO, 1991). Women in the developing countries bear about 200 times greater risk of dying from pregnancy related complications than those in the developed world (Mahler, 1987). Behind the death of each woman resulting from pregnancy complications there are many more women who suffer from serious, even long-term non-fatal pregnancy complications and in many situations without taking any appropriate measures for them. The earliest study (Datta *et al*, 1980), conducted on maternal morbidity in rural India, reported that for every maternal death 16.5 complications taking place related to pregnancy or puerperium. Based on this estimate Walsh *et al* (1989) estimated about 8 million pregnancy complications occur every year in the world. In 1993 Koblinsky *et al*, after evaluating several population-based studies in developing countries, assessed 40% of the pregnancies as complicated which revealed a several fold higher estimated total number of maternal morbidities than the earlier estimate, arising every year globally.

In Bangladesh since 1967-1968 several well designed maternal mortality studies (Chen *et al*, 1974; Khan *et al*, 1986; Alauddin, 1986; Koenig *et al*, 1988) have suggested that there has been a decline in maternal mortality in Bangladesh from 7.7 in 1967-1968 to 5.5 in 1976-1985 per 1,000 live births. In 1996 the maternal mortality ratio was estimated as 4.3 per 1,000 live births in Bangladesh (Islam and Hossain, 1997) which is still among the highest in the world. The severity of the problem of maternal morbidity in Bangladesh is also well conceived from these high maternal mortality ratios.

In 1992 a multinational collaborative retrospective study (Fortney and Smith 1996; Akhter *et al*, 1996) on maternal morbidity was conducted in four developing countries - Bangladesh, India, Egypt and Indonesia. We joined that study and

Correspondence: Mahbub-E-Elahi Khan Chowdhury, Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai 90110, Thailand.
E-mail: g4228005@maliwan.psu.ac.th

added a follow-up component with currently pregnant women to assess the level of their pregnancy-related complications throughout the course of pregnancy. Duration of identification of pregnant women from the community for this study was September-October 1992 and admission of subjects into the follow-up process was from October to December 1992.

The objectives of this study were (1) to estimate the occurrence of symptoms of various antepartum morbidities during different trimesters of pregnancy, (2) to assess the health burden of the pregnant women due to those morbidities, (3) to find out the care-seeking behavior of the pregnant women for their reported symptoms of morbidities among a representative sample of rural Bangladeshi population.

METHODS

Study design and sampling

This was a community-based cohort study. The sample was collected from rural areas of all the four old administrative divisions, Dhaka, Chittagong, Rajshahi and Khulan, of Bangladesh. In Bangladesh the administrative hierarchy is division, district, thana and union. We selected one district randomly from each division and from each of the selected districts one thana was selected randomly. From each of the selected thanas two unions were randomly selected. We included all the villages of the selected eight unions to identify married women who were within 24 weeks of gestation to enroll into this study.

Training of interviewers and quality control of data

A total of 12 female interviewers, 3 for each division, was recruited. A male supervisor supervised each of the teams. All of the interviewers and supervisors were graduates in social science or related subjects from local universities. A two-week intensive training course was given to them by the experienced researchers. In addition to training on basic interview techniques the training also emphasized maternal morbidity. The interviewers were also trained to minimize the inter-observer variation. Research physicians visited data collection spots at biweekly intervals and gave necessary guidelines to supervisors and interviewers to assure the quality of the collected data.

Identification of subjects and admission into the study

At first, the interviewers visited all the households in the selected unions to come up with a master list of currently pregnant women within 24 weeks of gestation among the currently married. The gestational age of the women were calculated by the interviewers from the first day of their last menstrual period. Later on they filled-out an admission form for each of the subjects and enrolled them into this study. At admission women were interviewed for their socio-demographic and reproductive characteristics. This form also included data for their morbidities such as hypertension, diabetes, pulmonary tuberculosis and jaundice during the current pregnancy.

Antenatal follow-up

The interviewers visited the subjects at home at every 4 weeks and obtained the symptoms of their current pregnancy-related morbidities. Women who reported symptoms of any morbidity were asked whether they had problems in their daily activities due to that morbidity and the types of problems they had. Women were also asked whether they sought care for their reported health problems and in case of seeking care, the person who provided care, was also recorded.

We defined the 1st trimester of pregnancy as completion of 14 weeks of gestation, the 2nd trimester as over 14 weeks to completion of 28 weeks of gestation, and the 3rd trimester as over 28 weeks of gestation. We defined bleeding as hemorrhage from the genital tract occurring after first trimester but before childbirth, fits and convulsions as a state in which the woman's body was affected by convulsion and eventually passed into coma. We defined excessive vomiting as persistent vomiting perceived serious by the pregnant woman, fever > 3 days as fever accompanying any infectious illness like burning urination or vaginal discharge, headache as intense pain felt deep in the skull. Symptomatic anemia was defined as pallor of conjunctivae and palm, palpitation was defined as awareness of the heart beat. Urinary problem was defined as irritation in the urinary tract or in mucous membrane of the genital tract during urination. Health burden was defined as constraints in daily activities of the pregnant women due to morbidities. Weights were applied to each report of morbidity according to intensity of constraints to estimate the summary measures of health burden. Constraints in activities were

categorized as follows - bed ridden for at least 3 days due to a morbidity and inability to do any work was defined as heavy health burden, daily activities hampered was defined as moderate burden, ability to continue daily activities despite difficulties as minimum burden, reported no problem in daily activities as no health burden. The mean intensity of burden for each episode of a morbidity was the weighted mean of burden over all the episodes of that morbidity. The population burden per 1,000 person months of a morbidity was the weighted mean of burden per 1,000 person months of observation of all the women in this study.

Statistical analysis

Descriptive analysis was used for percentages of symptoms of antepartum morbidities and percentages of types of burden in different trimesters of pregnancy. The health burden weighting scheme was as follows - bed-ridden and could not do any work:heavy=3, normal activities hampered : moderate=2, continued activities despite difficulties : minimum=1, no problem in activities : none=0.

RESULTS

Characteristics of sampled women (Table 1)

The sampled women had a low mean age at marriage (14.8 years) but at the time of interview (at mean age of 23 years) their parity was still relatively low (mean of 2.1). Their level of education and monthly per capita household expenditure were somewhat low. These characteristics were almost homogeneous except for parity, which was high in Chittagong (mean of 3.1) and low (mean of 1.4) in Rajshahi.

Dynamic cohort of the sampled women (Table 2)

Among 1,019 pregnant women 777 (76.3%) women were admitted into the study within 24 weeks of gestation as planned. The remaining 242 (23.7%) women entered after 24 weeks of gestation because there was a time lag between getting the master list and the time of first interview. During follow-up, some of the women were not at home and number of missed visits increased with increased gestation. Ten women were lost to follow-up - all of them due to a change of their usual residence. During the follow-up period, 37 (3.6%) reported having miscarriage and only 8 (0.8%) women admitted having induced abortion. We interviewed 964 women regarding their delivery. There was a total of 934 live births and 41 still births. We recorded 3 maternal deaths, occurring within 42 days of delivery, which gave an estimate of maternal mortality ratio of 3.2 per 1,000 live births.

Baseline morbidity from the admission records

Nearly 2% of the pregnant women reported having been diagnosed of having hypertension, more than 1% diabetes, nearly 4% jaundice, less than 1% (0.3%) pulmonary tuberculosis.

Occurrence of symptoms of antepartum morbidities (Table 3)

The percentages of visits in any trimester at which symptoms of morbidities were reported ranged

Table 1
Characteristics of the sampled women in 4 administrative divisions of Bangladesh

Characteristics	Dhaka n=214	Chittagong n=262	Rajshahi n=274	Khulna n=269	Total n=1,019
Mean age of respondents in years (SD)	23.8 (5.7)	25.6 (6.7)	21.4 (5.4)	21.5 (5.3)	23.0 (6.1)
Mean age at marriage in years (SD)	15.7 (2.3)	14.8 (2.2)	14.7 (2.1)	14.1 (2.5)	14.8 (2.3)
Mean no. of parity (SD)	2.3 (2.3)	3.1 (2.6)	1.4 (1.7)	1.6 (1.9)	2.1 (2.2)
Percentage never attended school	55.1	46.9	56.2	62.5	55.3
Median monthly per capita expenditure (in Taka)	333.00	364.00	384.00	375.00	333.00

Table 2
Dynamic cohort of the sampled women.

Gestational weeks	Number of women						No. of antepartum visits
	Admitted	Miscarriage	Induced abortion	Delivered	Lost to follow-up	Continued pregnancy	
05-08	7	3	0	-	-	4	4
09-12	55	2	3	-	-	54	54
13-16	170	6	4	-	1	213	202
17-20	309	10	0	-	2	510	490
21-24	236	4	0	-	1	741	680
25-28	173	8	1	3	2	900	802
29-32	59	4	0	23	3	929	793
33-36	8	-		98	0	839	578
37-40	2	-		455	1	385	184
40+	-	-		385	-	-	25

Table 3
Symptoms of various antepartum morbidities among rural Bangladeshi women.

Symptoms of morbidity*	%(number) of visits in trimesters				%(number) of women
	1st n=134	2nd n=2098	3rd n=1580	All n=3812	n=1019
Bleeding		0.3 (6)	0.3 (4)	0.3 (10)	1.0 (10)
Fits and convulsions		0.5 (10)	1.1 (17)	0.7 (27)	2.3 (23)
Swelling of hands and legs		2.0 (43)	5.4 (86)	3.4 (129)	9.1 (93)
Excessive vomiting	4.5 (6)	1.4 (30)	1.2 (19)	1.4 (55)	4.4 (45)
Fever > 3 days	5.3 (7)	3.6 (75)	4.5 (71)	4.0 (153)	12.4 (126)
Headache	30.3 (40)	22.6 (474)	19.6 (309)	21.6 (823)	43.7 (445)
Urinary problem	23.5 (31)	26.3 (552)	27.5 (434)	26.8 (1,021)	48.2 (491)
Palpitation	37.9 (50)	44.9 (942)	49.5 (782)	46.5 (1,774)	69.0 (703)
Symptomatic anemia	65.2 (86)	74.5 (1,564)	84.4 (1,334)	78.3 (2,984)	91.7 (934)
None of the above	25.0 (33)	18.4 (386)	10.3 (162)	15.2 (581)	6.1 (62)

*Multiple responses.

Table 4
Percentage (number) of visits in each trimester in which women in rural Bangladesh experienced types of health burden for symptoms of any antepartum morbidity.

Type of burden	%(number) of visits in trimesters			
	1st n=101	2nd n=1,712	3rd n=1,418	All n=3,231
Bedridden - could not do any work	1.0 (1)	1.1 (19)	1.3 (18)	1.2 (38)
Normal activities hampered	11.9 (12)	7.2 (123)	7.6 (108)	7.5 (243)
Continued activity with difficulty	62.4 (63)	61.2 (1,048)	58.6 (831)	60.1 (1,942)
No problem in activity	24.8 (25)	30.5 (522)	32.5 (461)	31.2 (1,008)

Table 5
Health burden for different antepartum morbidities of the pregnant women in rural Bangladesh.

Symptoms of morbidity	No. of visits with types of burden				Weighted burden	
	Heavy n=38	Moderate n=243	Minimum n=1,942	None n=1,008	Mean intensity of burden	Population burden /1,000 person months
Bleeding	1	1	6	2	1.10	3.63
Fits and convulsions	8	8	10	1	1.85	16.52
Swelling of hands and legs	7	13	98	11	1.12	47.90
Excessive vomiting	3	10	32	10	1.11	20.15
Fever > 3 days	4	22	87	40	0.93	47.24
Headache	16	97	562	148	0.98	265.61
Urinary problems	11	105	616	289	0.84	283.78
Palpitation	20	169	1,092	493	0.84	492.24
Symptomatic anemia	29	216	1,810	929	0.78	769.41

Total time observed 1,019 women for antepartum morbidities = 3,027 person months.

form 0.3% for bleeding to 78.3% for symptomatic anemia. The majority of the symptoms were more common in the second and third trimesters except excessive vomiting, headache and fever. The percentages of women who reported symptoms of morbidity at any time during pregnancy were 1% for antepartum bleeding and above 2% for fits and convulsions.

Health burden due to morbidity (Tables 4 and 5)

Most of symptoms of morbidity led to difficulty in daily activity in one way or another. Symptoms in the 1st trimester gave more health burden (75.2%) than those in the 2nd (69.5%) and 3rd (67.5%) trimesters. There was a inverse relationship between mean intensity of burden and population burden for most of the morbidities. For example, for fits and convulsions mean intensity of burden per episode of this morbidity was 1.85 and for symptomatic anemia it was only 0.78 whereas the population burden per 1,000 person months for fits and convulsions was about 4 but that for anemia was more than 700.

Care-seeking pattern for symptoms of morbidities (Table 6)

The percentage of visits in which women did not seek care from any provider for experiencing symptoms of different antepartum morbidities ranged from 25.9% for fits and convulsions to 87.4% for symptomatic anemia. For any episodes of convulsion and bleeding more than half of the attacks resulted in care seeking, in which majority of the care providers were either quack doctors or traditional healers.

DISCUSSION

Women in rural Bangladesh in our study were at a low socio-economic status and the symptoms of antepartum morbidities were very common. Health burden associated with their reported symptoms of morbidities was very heavy. Almost all of the women who had any symptom of morbidity had difficulty in one way or another in doing their daily activities. On the other hand, care-seeking behavior of the women having those symptoms was very poor. For most morbidities the majority of the times women did not seek care from any provider.

The number of women lost to follow-up from this study was quite low. The percentage of visits at which the woman was absent was low in 1st and 2nd trimesters but increased in the 3rd trimester because by tradition rural Bangladeshi women prefer to go to their father's home to deliver the baby and return to their husband's home afterwards.

Our finding of early marriage at the mean age of 14.2 years is consistent with the national statistic of median age at first marriage of the rural Bangladeshi women 14.0 years (Mitra *et al*, 1997). The majority of our study subjects (55.3%) never

Table 6
Percentage (number) of visits in which women in rural Bangladesh sought care from various types of providers for symptoms of different antepartum morbidities.

Care provider	Fits and convulsions n=27	Bleeding n=10	Fever>3days n=153	Excessive vomiting n=55	Headache n=823	Swelling of hands and face n=129	Symptomatic anemia n=2,984	Urinary problem n=1,021	Palpitation n=1,774
Physician	33.3 (9)	10.0 (1)	7.8 (12)	12.7 (7)	6.8 (56)	7.8 (10)	7.9 (235)	4.4 (45)	5.9 (105)
FWV/FWA/TBA/paramedic*	-	-	1.3 (2)	1.8 (1)	0.1 (1)	0.8 (1)	1.2 (36)	0.8 (8)	0.5 (9)
Homeopath	3.7 (1)	-	4.6 (7)	1.8 (1)	2.3 (19)	2.3 (3)	0.8 (24)	2.3 (23)	1.0 (18)
Quack*	33.3 (9)	10.0 (1)	19.0 (29)	14.5 (8)	13.5 (111)	6.2 (8)	7.6 (226)	5.9 (60)	4.3 (76)
Traditional*	3.7 (1)	30.0 (3)	1.3 (2)	-	2.4 (20)	0.8 (1)	1.0 (29)	2.4 (25)	0.8 (15)
Relative/neighbor	-	-	-	1.8 (1)	0.2 (2)	3.1 (4)	0.7 (20)	1.6 (16)	-
Sought no care	25.9 (7)	50.0 (5)	66.0 (101)	67.3 (37)	74.6 (614)	79.1 (102)	80.9 (2,412)	82.7 (844)	87.4 (1,551)

*FWV/FWA/TBA/paramedic - Family Welfare Visitor/Family Welfare Assistant/Trained Traditional Birth Attendant/Medical Assistant.
*Quack - Unqualified village doctor, *Traditional - Herbal medicine from traditional provider called Kabiraj

attended school, which is also consistent with the national data (of 57.1 %) of ever married rural women never having attended school (Mitra *et al*, 1997). This early marriage combined with the hardship of poverty and low education may easily make these women vulnerable to various complications during pregnancy. Differences in parity between two sampled areas, Chittagong and Rajshahi in our study is ineversely related with the contraceptive prevalence rate (CPR) in those two divisions of Bangladesh. CPR is low in Chittagong and high in Rajshahi; for example in 1993-1994 CPR in these two divisions were 29.3% and 54.8%, respectively (Mitra *et al*, 1997).

We can compare the 3.7% miscarriage among 1,019 women of our study with that of 5.0% among 33,473 pregnancies during 1982-1991 in Matlab, Bangladesh (Ahmed *et al*, 1998). However, our finding for induced abortion ratio of 8.6/1,000 live births was much lower than 20.0/1,000 live births in Matlab of the same study. The reasonable explanation for underestimation of our these two statistics in our study is that majority of our sampled women entered into the follow-up process after the first trimester. The still birth ratio of 43.9/1,000 live births of our study is somewhat higher than the estimate of 35.3/1,000 live births of the same study in Matlab, which is an intervention area of ICDDR,B (International Center for Diarrheal Diseases and Research, Bangladesh) for family planning and maternal and child health care. Our estimated maternal mortality ratio of 3.2/1,000 live births can be compared with the 1996 estimate of 4.3/1,000 live births (Islam and Hossain, 1997) and the slight difference may be due to chance for observation of rare events like maternal death.

Among baseline morbidities our findings on jaundice (4.0%) and pulmonary tuberculosis (0.3%) can be compared with 5.9 and 0.3%, respectively, of the retrospective study (Fortney and Smith, 1996). The same study found 3.6% of the women had hypertension during pregnancy whereas nearly 2% of our subjects reported diagnosed hypertension at the time of admission.

One percent of women in our study experienced antepartum bleeding either in 2nd or 3rd trimester. Among the other studies conducted in Bangladesh the Maternal Care Project in Matlab also found 1% antepartum hemorrhage (Stewart and Maxine, 1991). The retrospective study conducted among Indian women (Bhatia, 1995) found 0.9% women experiencing antepartum bleeding.

Fits and convulsions were reported 4 times more commonly in the 3rd trimester than in the 2nd trimester. More than 2% of the pregnant women in our study had fits and convulsions which can be compared with the findings of 3% fits/convulsion (Fortney and Smith, 1996) and 6% severe eclampsia found in Maternal Care Project in Matlab (Stewart *et al*, 1991).

The fact that fever >3 days and urinary problems were common may be due to high incidence of urinary tract infections. This suggests a need to study to confirm the clinical nature of these two problems. Palpitation and symptomatic anemia (as judged by the trained home visitors) were very common and increased with increased gestational age. Iron deficiency anemia is known to be very common in rural areas of South Asian countries due to poverty and undernourishment (Stewart and Maxine, 1991; Krishnaswami, 1998; DeMaeyer *et al*, 1985). This may be the explanation for the commonness of these symptoms in our population.

Despite the problem of clinical uncertainties over the precise medical diagnosis of the symptoms of antepartum morbidities, our study has shed some light on the extent of the burden on the pregnant women's daily life activities. The measures of burden are likely to be valid measures of the problem in such an area where medical facilities are not available. Looking at the seriousness of the symptoms alone might not be enough to reflect the magnitude of burden. Our study show that mild symptoms (such as palpitation and urinary problems) are very common. Each of these brings a little bit of burden every day but its commonness and long lasting nature lead to a large magnitude of population burden compared to those more serious symptoms such as bleeding and fits.

Poor care-seeking behavior has been observed among the pregnant women for their antepartum morbidities. Except for fits and convulsions, the majority of the times women sought no care for their health problems. For about three fourth of the episodes of having fits and convulsions, women sought care from a provider. If we consider physicians and family planning workers like family welfare visitors (FWV) and family welfare assistant (FWA) including trained traditional birth attendants (TTBA) and paramedics as trained providers and the rest as untrained providers for management of pregnancy complications, then the village women had a tendency to visit untrained providers, especially the quack, more frequently than their trained counterparts.

CONCLUSION

Since the symptoms of antepartum morbidities were very commonly reported by rural Bangladeshi women there is a need for further investigation to explore the underlying reasons for this immense problem. Further studies with appropriate clinical facilities are recommended. Care-seeking behavior of the pregnant women of rural Bangladesh was very poor whereas trained health care providers like FWVs, FWAs, TTBAs and paramedics were grossly underutilized. There is a need to conduct a research project to find out the reasons for this problem. Further understanding of the nature of these morbidities and the health care-seeking behavior will assist the implementation of a more appropriate prevention and treatment program for poor pregnant women of rural Bangladesh

ACKNOWLEDGEMENTS

This paper is a part of first author's thesis to fulfil the requirements for his PhD in Epidemiology under BIRPERHT/Ford Foundation fellowship at Prince of Songkla University (PSU), Thailand. The Epidemiology Unit at PSU is partially supported by Thailand Research Fund under the grant for Senior Research Scholar of Dr Virasakdi Chongsuvivatwong. We express our gratitude to BIRPERHT for providing the data of the "Maternal Morbidity Study" for this publication. We acknowledge the financial support for this study from Ford Foundation. We would like to thank all the researchers, interviewers and editors for the sincere and enduring efforts they provided for this study. The authors gratefully acknowledge Dr Alan Geater for his reviewing the manuscript and suggestions.

REFERENCES

Ahmed MK, Rahman M, Ginneken J. Induced abortion in Matlab, Bangladesh: Trends and determinants. *Int Fam Plann Perspect* 1998; 24: 128-32.

Akhter HH, Chowdhury MEK, Sen A. A cross-sectional study on maternal morbidity in Bangladesh. Dhaka: Bangladesh Institute of Research for Promotion of Essential and Reproductive Health and Technologies (BIRPERHT), 1996.

Alauddin M. Maternal mortality in rural Bangladesh: The Tangail district. *Stud Fam Plann* 1986; 17: 13-21.

Bhatia JC. Levels and determinants of maternal morbidity: results from a community-based study in southern India. *Int J Gynecol Obstet* 1995; 50(suppl 2): 153-63.

Chen LC, Gesche MC, Ahmed S, Chowdhury AI, Mosley WH. Maternal mortality in rural Bangladesh. *Stud Fam Plann* 1974; 5: 334-41.

Dutta KK, Sharma RS, Razack PM, Ghosh TK, Arora RR. Morbidity pattern amongst rural pregnant women in Alwar, Rajasthan - a cohort study. *Health Popul Perspect Issues* 1980; 3: 282-92.

DeMaeyer E, Adiels-Tegman M. The prevalence of anemia in the World. *World Health Stat Q* 1985; 38: 302-16.

Fortney JA, Smith JB, eds. The base of the Iceberg : Prevalence and perceptions of maternal morbidity in four developing countries. The Maternal Morbidity Network. Research Triangle Park, North Carolina: Maternal and Neonatal Health Center, Family Health International (FHI), 1996.

Islam W, Hossain MS. Reproductive health status in Bangladesh. Dhaka: Bangladesh Bureau of Statistics (BBS), 1997; (monograph series 06).

Khan AR, Jahan FA, Begum SF. Maternal mortality in rural Bangladesh: The Jamalpur district. *Stud Fam Plann* 1986; 17: 7-12.

Koblinsky MA, Campbell OM, Harlow SD. Mother and more: A broader perspective on women's health. In: Koblinsky M, Timyan J, Gay J, eds. The Health of Women: A Global Perspective. Oxford: Westview Press, 1993: 33-62.

Koenig M, Chowdhury AI, Fauveau V, Chakraborty J. Maternal mortality in Matlab, Bangladesh: 1976-85. *Stud Fam Plann* 1988; 19: 69-80.

Krishnaswami K. Country profile : India. Nutritional disorders - old and changing. *Lancet* 1988; 351: 1268-9.

Mahler H. The safe motherhood initiative : A call to action. *Lancet* 1987; 1: 668-70.

Mitra SN, Al-Sabir A, Cross AR, Jamil Kanta. Bangladesh Demographic and Health Survey (BDHS) 1996-97. Dhaka: National Institute of Population Research and Training (NIPORT), Mitra and Associates and Macro International Inc, 1997.

Stewart Kate, Maxine Whittaker. Methodological issues in defining female morbidity : A case study from the Maternity Care Project, Matlab, Bangladesh Paper presented at the 18th Annual NCIH International Health Conference, Arlington, Va. 1991.

Walsh JA, Nashak CM, Measham AR, Gertler PJ. Maternal and Perinatal Health Problems. In: Jamison DT, Mosley WH, eds. Evolving Health Sector Priorities in Developing Countries. Washington, DC: The World Bank, 1989.

World Health Organization (WHO). Maternal mortality ratios and rates: A tabulation of available information. 3rd ed. *WHO/MCH/MSM/91.6*, 1991.

FOLLOW UP OF WATER USE IN A TIN MINING AREA AFFECTED WITH ARSENIC POISONING

Virasakdi Chongsuvivatwong, Apiradee Lim, Mafaosis Dueravee, Alan Geater, Skulrat Ritsamitchai and Shoko Oshikawa

Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai 90112, Thailand

Abstract. Ron Phibun district in southern Thailand has been known as an endemic area for arsenic contamination. The government has been trying to improve the situation by encouraging the use of rainwater and piped water. This study aimed to document the change of water use and to identify factors associated with safe water use in 1997 compared to that in 1994. Home visits and face-to-face questionnaire interviews were undertaken. Information on water use for drinking, cooking, washing food and washing utensils in 1994 and 1997 was obtained. Among 3,849 households from which data could be obtained (estimated 79% of total households), the percentages of using safe water (including water from bottled rain water, piped and artesian well water) for drinking and cooking rose from 72.5 and 57.9 in 1994 to 93.6 and 80.9 in 1997, respectively. The percentages for washing foods and for washing utensils rose from 28.6 and 20.5 to 59.1 and 53.8, respectively. In 1997, percentage of households using piped water for drinking and cooking was still low (3.6 and 12.3) compared to those using piped water for washing food and utensils (39.1 and 43.6). Multivariate analysis shows that independent factors of the household predicting safe water use are: high arsenic area, near main road and having piped water installed. The influence of these factors (as judged by the level of odds ratio) operates more or less equally on water use for all purposes, except that installation of piped water has more influence on washing water than drinking and cooking water. We conclude that safe water supply in the area is still inadequate. Even if piped water is installed, it is often not used for drinking and cooking. The reasons for not using piped water for drinking and cooking need to be identified.

INTRODUCTION

Problems of arsenic contamination in water leading to arsenosis are not uncommon. Endemic arsenosis has been documented in Japan (Tsuda *et al*, 1995), Taiwan (Smith *et al*, 1992), India (Chakraborty and Saha, 1987; Guha Mazumder *et al*, 1988; 1992; Saha, 1995), Bangladesh (Nickson *et al*, 1998; Tondel, 1998; Tondel *et al*, 1999), China (Wu, 1993; Zhang and Chen, 1997), Mongolia (Luo *et al*, 1995), Mexico (Cebrian *et al*, 1983), Argentina (Astofi *et al*, 1981) and Chile (Borgono *et al*, 1977). This problem also occurs in Thailand.

In addition to development of skin lesions and skin cancer, arsenic contamination has been demonstrated to be associated with increased risk for hypertension (Rahman *et al*, 1999), diabetes mellitus (Rahman *et al*, 1998) and internal cancers (Ferreccio *et al*, 1998; Karagas *et al*, 1998). There is a need to improve the situation by the supply of safe water supply to the affected communities. The current article illustrates an evaluation of water supply program in a contaminated area in Thailand.

In 1987, more than 1,000 cases of skin arsenosis were reported from Ron Phibun Sub-district, Nakhon Si Thamarat Province, southern Thailand (Arrykul *et al*, 1996). This area has been the site of tin mining activities for almost a century. Although most of the mines ceased operation in the late 1980s, the release of arsenopyrite (FeAsS) from the tin ore has left a legacy of extensive arsenic contamination in many areas of Ron Phibun District. Groundwater and many shallow wells are contaminated by arsenic (Boriboon, 1996) above the acceptable level for drinking water set by WHO (0.01 ppm) (WHO, 1993) expecially in villages 2, 12 and 13 (Paijitprapaporn, 1997).

Since arsenic contamination of shallow well water, surface water and soil was detected in 1987, various governmental departments have tried to solve this problem. Water jars for storing rainwater were distributed and piped water systems were expanded. However, the effect of such intervention has never been evaluated and changing water use among the villagers had not been investigated. This study aimed to investigate the change of water use between 1994 and 1997 and to identify

Correspondence: Virasakdi Chongsuvivatwong, Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai 90112, Thailand.
E-mail: cvirasak@ratree.psu.ac.th

the factors associated with safe water use in 1997. Experience from this program may be useful in further planning of intervention in this community as well as in other affected areas in developing countries.

MATERIAL AND METHODS

Study sites

The study area consists of 16 villages with a population of approximately 15,095 and 4,900 households three of these villages (2, 12 and 13) are known to have high prevalence of shallow well water found contaminated with arsenic (Table 1).

Data collection

Following a survey in April-November 1996, a map of each village was drawn showing location of households, roads, key sites and distribution of pipelines of water supply. Home visits were made between April and May 1997 by a group of researchers and medical students. A structured questionnaire was tested and used to collect data by face-to-face interview with the head of the family or other available member. Variables collected consisted of household and family head characteristics, and water use (including source of water for drinking, cooking, washing food and washing utensils) in 1994 and 1997.

Data management and statistical analysis

Data collected with the questionnaire were computerized using Epi Info software. The map was digitized using Arc-Info software for geographical query. Stata version 6 program package was used for statistical analysis. Descriptive statistics were computed. Types of water use were bottled, piped, artesian well, rain, shallow well, and stream water. The first four were classified as safe and the latter two as unsafe. Factors associated with safe water use in 1997 were identified using logistic regression.

RESULTS

Data were obtained from 3,849 households out of 4,900 (an availability proportion of 79%). Incomplete data collection was due to the fact that many houses were closed during the home visit. This was common in the suburban area of village 12 where people went out to work in the city. Breakdown by village is shown in Table 1.

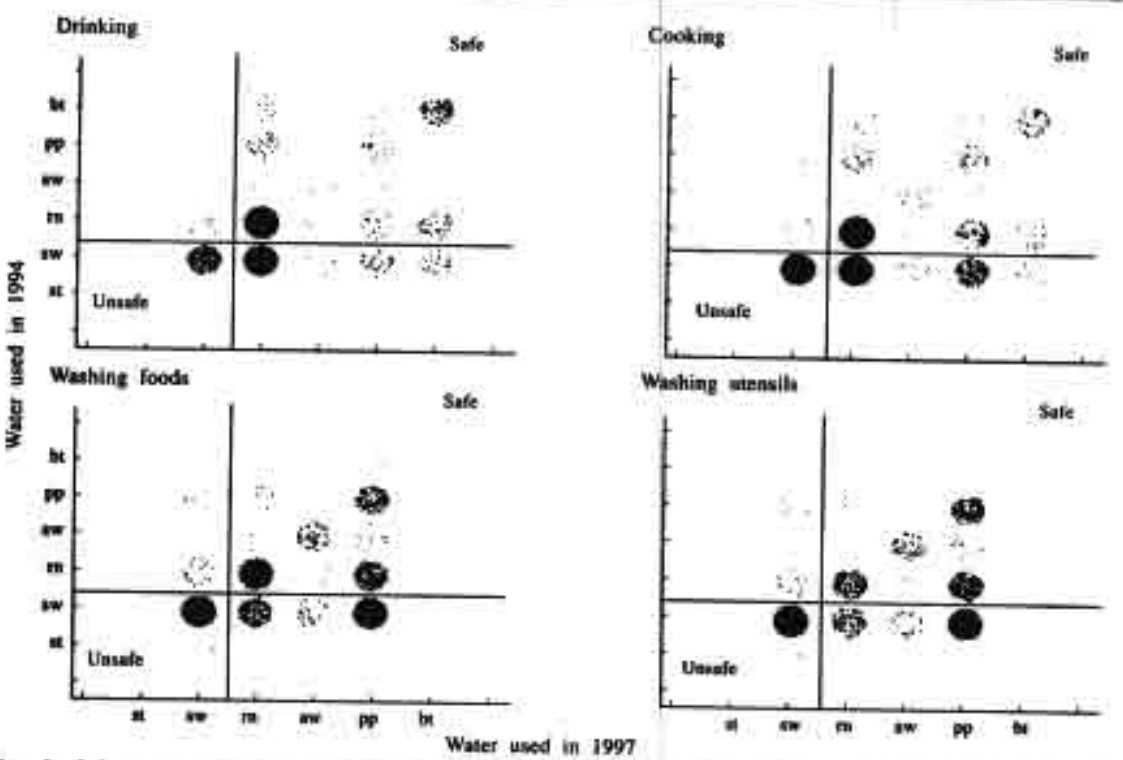
Table 1
Characteristics of study villages sorted by proportion of shallow wells with arsenic contaminated water.

Village	Total number of visited households	Number of household with data	Population in households with data	% of wells with As>0.05 (ppm[a])
2	378	329	1,145	35.4
12	672	404	1,502	30.4
13	525	290	1,099	25.5
5	282	246	1,064	10.0
1	131	116	497	9.2
8	309	240	981	4.5
10	306	229	924	4.2
14	172	144	690	2.6
9	214	193	859	2.2
3	330	272	1,013	1.6
4	279	256	1,260	1.5
7	643	484	1,514	0.9
6	231	207	965	0.6
11	147	133	497	0.0
15[b]	80	65	242	-
16[b]	201	177	834	-
Total	4,900	3,849	15,095	

[a]Reported in 1994 (Choprapawon, 1994).
[b]Village 15 was previously part of village 1 and village 16 part of village 9.

Table 2
Water used in 1994 and 1997.

Type of water	Number (%) of household using water for			
	Drinking	Cooking	Washing food	Washing utensils
Water used in 1994	n = 3,255	n = 3,209	n = 3,147	n = 3,134
Bottled	118 (3.6)	61 (1.9)	2 (0.1)	1 (0.0)
Piped	47 (1.4)	66 (2.1)	174 (5.5)	186 (5.9)
Artesian well	6 (0.2)	23 (0.7)	79 (2.5)	91 (2.9)
Rain	2,189 (67.3)	1,706 (53.2)	646 (20.5)	365 (11.7)
Shallow well	895 (27.5)	1,353 (42.2)	2,245 (71.3)	2,490 (79.5)
Stream			1 (0.0)	1 (0.0)
Water used in 1997	n = 3,785	n = 3,783	n = 3,783	n = 3,784
Bottled	257 (6.8)	136 (3.6)	3 (0.1)	
Piped	137 (3.6)	466 (12.3	1,479 (39.1)	1,650 (43.6)
Artesian well	9 (0.2)	34 (0.9)	108 (2.9)	114 (3.0)
Rain	3,143 (83.0)	2,424 (64.1)	643 (17.0)	275 (7.3)
Shallow well	239 (6.3)	722 (19.1)	1,550 (41.0)	1,745 (46.1)
Stream		1 (0.0)		



bt = bottled water; pp = piped water; aw = artesian well water; rn = rainwater; sw = shallow well water; st = stream water

Fig 1–Scatter plot of type water used in 1994 against type of water used in 1997.

Table 2 shows that the situation of water use in 1997 was improved compared to that in 1994. Of various types of water source for drinking and cooking, there was a noticeable increase in rainwater and slight absolute increase in percentage in bottled and piped water. Shallow well water use for these purposes decreased. For washing purposes, use of piped water was remarkably increased whereas use of rainwater was decreased. For more detail, a scatter plot is shown in Fig 1.

Table 3
Factor associated with using safe water in 1997 for various purposes.

Household Characteristics	No. (%) of households using safe water use			
	Drinking (n=3,785)	Cooking (n=3,783)	Washing food (n=3,783)	Washing utensils (n=3,784)
High contamination area				
No (ref)	2,540 (92.0)	2,080 (75.3)	1,415 (51.3)	1,277 (46.2)
Yes	1,006 (98.3)	980 (95.9)	818 (80.0)	762 (74.5)
cOR (95% CI)	5.2 (3.1-8.5)	7.6 (5.5-10.5)	3.8 (3.2-4.5)	3.4 (2.9-4.0)
aOR (95%CI)	3.1 (1.9-5.2)	4.1 (2.9-5.8)	1.9 (1.5-2.4)	1.4 (1.1-1.9)
Distance from main road (<200 m)				
No (ref)	2,260 (92.7)	1,857 (76.1)	1,285 (52.7)	1,151 (47.2)
Yes	1,286 (95.5)	1,203 (89.5)	948 (70.5)	888 (66.0)
cOR (95%CI)	1.7 (1.3-2.3)	2.7 (2.2-3.3)	2.1 (1.9-2.5)	2.2 (1.9-2.5)
aOR (95%CI)	1.1 (0.8-1.5)	1.7 (1.3-2.1)	1.4 (1.1-1.7)	1.5 (1.2-1.9)
Access to piped water				
No (ref)	1,869 (89.8)	1,409 (67.7)	577 (27.7)	376 (18.1)
Yes	1,677 (98.5)	1,651 (97.1)	1,656 (97.4)	1,663 (97.7)
cOR (95% CI)	7.4 (4.9-11.1)	15.8 (11.7-21.2)	96.0 (70.3-131)	194 (138-271)
aOR (95% CI)	5.8 (3.8-8.8)	12.6 (9.3-17.0)	89.4 (64.8-123)	183 (129-158)
Education of family head				
≤ Primary (ref)	835 (94.8)	761 (86.4)	574 (65.2)	527 (59.8)
> Primary	2,657 (93.3	2,255 (79.2)	1,629 (57.2)	1,482 (52.1)
cOR (95% CI)	1.3 (0.9-1.8)	1.7 (1.3-2.1)	1.4 (1.2-1.6)	1.4 (1.2-1.6)
aOR (95% CI)	1.0 (0.7-1.3)	0.8 (0.6-1.0)	0.9 (0.7-1.1)	0.9 (0.7-1.2)

Fig 1 is a "jittered" scatter plot, in which source of water in 1994 (Y axis) is plotted against that in 1997 (X axis). The vertical lines divide water use in 1997 into "safe"on the left side and "unsafe" on the right. Horizontal lines divide water use in 1994 into "safe" above the line and "unsafe" underneath. Dots in the central diagonal represent households not changing water source. As expected, the major changes in drinking and cooking (upper two graphs) were from shallow well to rain water, whereas changes in washing (lower two graphs) were from various types to piped water.

Fig 2 shows percent of households using safe water in 1994 (base part of the arrows) and 1997 (point of the arrow) of villages sorted into ascending order along the X-axis by percent of safe water use in 1994. Bold arrows indicate villages with high levels of contamination. Drinking water has the highest overall percentages of being safe, both in the baseline (1994) and after the change (1997). Water for cooking and washing was not as safe. Most households in the heavily contaminated villages with high contamination turned to use safe water for all purposes in 1997.

Table 3 displays the percentage of safe water use for different purposes broken down by various independent variables. Living in an area with high level of arsenic contamination, living closer to main road (except for drinking purpose), and having access to piped water are positively associated with use of safe water. However, it should be noticed that among these variables, access to piped water is the most powerful determinant as the odds ratios are very high. Again, it is shown that the effect of this variable is stronger for washing purposes than for drinking or cooking.

DISCUSSION

It is clear that water use pattern in the study area has improved as more households turned to use rainwater and piped water, which has been shown to be relatively arsenic-free (Oshikawa, 1997). However, there are still some important pitfalls in the program. First, at the time of evaluation, there

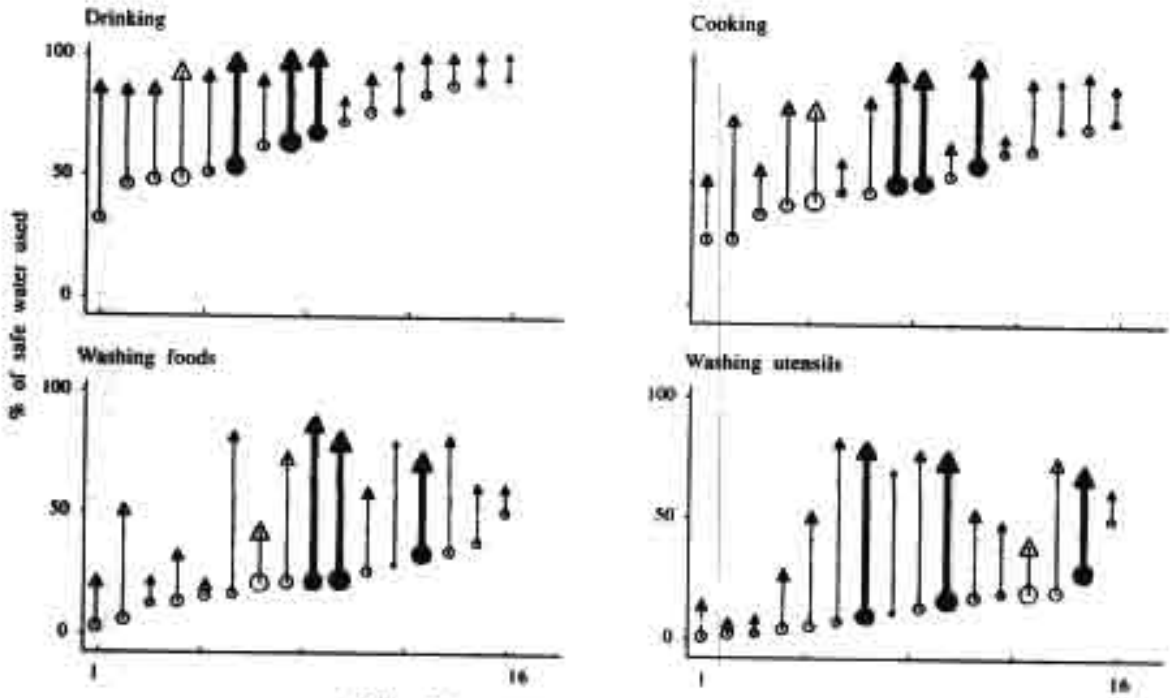


Fig 2–Sorted percent of households using safe water for drinking, cooking, washing food and washing utensils in 1994 in each of 16 villages.

were still 6 to 19% of the population using unsafe water for drinking and cooking and over 40% using it for washing food and utensils. Moreover, piped water was not popular for drinking and cooking but was used for mainly for washing. Thus the investment on provision of piped water did not reach its goal.

The use of rainwater should be encouraged with proper instruction regarding the collection process. Dusts in this area may contain arsenic and it can be accumulated on the roof of the house. The rainwater jar provided by the government was also difficult to clean (Adjimangkul, 1992). Therefore, it is strongly advised that rainwater should be collected only in high rainy season. One should let initial heavy rains wash away the dusts on the roof and gutters several days before the rainwater is collected.

The availability proportion in our study was 79%. We still lack data from more than one-fifth of the total households. The non-responders tended to go to work in nearby cities where arsenic contamination was not shown to be an important problem. Their lifestyles are likely to be more urbanized than the responder households. As our multivariate analysis demonstrated that households with higher level of urbanization (nearer to main road and access to piped water) are more likely to use safe water, the non-responders are therefore likely to have less serious problems.

In conclusions, the situation of water supply has been improved but still requires further effort to increase the coverage. Research on sociological aspects to explain low level of use of piped water for drinking and cooking should be initiated.

ACKNOWLEDGEMENTS

This study was supported by Thailand Research Fund.

REFERENCES

Adjimangkul S. Mothers' beliefs concerning water consumption and arsenic poisoning of children under 5 years of age: A community study in Nakhon Si Thammarat Province. MSc thesis. Bangkok: Mahidol University 1992.

Arrykul S, Kooptamon K, Wittayawarawat W. Contamination of arsenic, cadmium, and lead in Pakpanang river basin, Nakhon Si Thammarat, Thailand. Progress in research and control of chronic arsenic poisoning in Ron Phibun Subdistrict, Ron Phibun District, Nakhon Si Thammarat Province. Epidemiology Division, Ministry of Public Health: 1996: 45-54.

Astofi E, Maccagno A, Garcia Fernandez JC, Uaccaro R, Stimola R. Relational between arsenic in drinking water and skin cancer. *Rev Med Chi* 1981; 3: 133-43.

Borgono JM, Vicent P, Venturino H, Infante A. Arsenic in the drinking water of the city of Antofagasta: epidemiological and clinical study before and after the installation of a treatment plant. *Environ Health Perspect* 1977; 19: 103-5.

Boriboon P. Study of arsenic intake in daily consumption. Progress in research and control of chronic arsenic poisoning in Ron Phibun Subdistrict, Ron Phibun District, Nakhon Si Thammarat Province. Epidemiology Division. Ministry of Public Health; 1996: 31-44.

Cebrian ME, Albores A, Aguilar M, Blakely E. Chronic arsenic poisoning in the North of Mexico. *Human Toxicol* 1983; 2: 121-33.

Chakraborty AK, Saha KC. Arsenical dermatosis from tubewell water in West Bengal. *Indian J Med Res* 1987; 85: 326-34.

Choprapawon C. Arsenic problem in Ron Phibun District, Nakhon Si Thammarat Province. Thailand Health Research Institute, 1994.

Perreccio C, Gonzalez Psych C, Milosavjlevic Stat V, Marshall Gredis G, Sancha AM. Lung cancer and arsenic exposure in drinking water: a case-control study in northern Chile. (ad nude lublica) 1998; 14 (suppl 3): 193-8.

Guha Mazumder DN, Chakraborty AK, Ghose A, *et al*. Chronic arsenic toxicity from drinking tubewell water in rural West Bengal. *Bull WHO* 1988; 66: 499-506.

Guha Mazumder DN, Das Gupta J, Chakraborty AK, Chatterjee A, Das D, Chakraborti D. Environmental pollution and chronic arsenicosis in South Calcutta. *Bull WHO* 1992; 70: 481-5.

Karagas MR, Tosteson TD, Blum J, Morris JS, Baron JA, Klaue B. Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a US population. *Environ Health Perspect* 1998; 106 (suppl 4): 1047-50.

Luo FJ, Luo ZD, Ma L. A study on the relationship between drinking water with high arsenic content and incidence of malignant tumour in Heihe Village, western part of Hubehot, Inner Mongolia. *Chung Hua Liu Hsing Ping Hsueh Tsa Chih* 1995; 16: 289-91.

Nickson R, McArthur J, Burgess W, Ahmed KM, Ravenscroft P, Rahman M. Arsenic poisoning of Bangladesh groundwater. *Nature* 1998; 395-8.

Oshikawa S. Re-examination of a cohort of subjects with arsenical skin lesions ten years after inceptions Songkhla: Master's Degree thesis, Prince of Songkla University; 1997.

Paijitprapaporn A. Situation regarding the problem and remedy of arsenic distribution in Ronpiboon District, Nakorn Sri Thammarat Province. Problem-solving regarding the distribution of arsenic at Ronpiboon District, Nakorn Sri Thammart Province. Department of Mineral Resources; 1997: 1-24.

Rahman M, Tondel M, Ahmad SA, Axelson O. Diabetes mellitus associated with arsenic exposure in Bangladesh. *Am J Epidemiol* 1998; 148: 198-203.

Rahman M, Tondel M, Ahmad SA, Chowdhury IA, Faruquee MH, Axelson O. Hypertension and arsenic exposure in Bangladesh. *Hypertension* 1999; 33: 74-8.

Saha KC. Chronic arsenical dermatoses from tube-well water in West Bengal during 1983-87. *Ind J Dermatiol* 1995; 40: 1-12.

Smith AH, Hopenhayn Rich C, Bates MN, *et al*. Cancer risks from arsenic in drinking water. *Environ Health Perspec* 1992; 97: 259-67.

Tondel M. Arsenic poisoning in Bangladesh–an environmental disaster threatens millions of people. *Lakartidningen* 1998; 95: 3075-8.

Tondel M, Rahman M, Magnuson A, Chowdhury IA, Faruquee MH, Ahmad SA. The relationship of arsenic levels in drinking water and the prevalence rate of skin lesions in Bangladesh. *Environ Health Perspect* 1999; 107: 727-9.

Tsuda T, Babazono A Yamamoto E. *et al*. Ingested arsenic and internal cancer: A historical cohort study followed for 33 years. *Am J Epidemiol* 1995; 141: 198-209.

WHO. Guidelines for Drinking Water Quality, 2nd ed. Vol 1. Recommendations. Geneva: World Health Organization, 1993.

Wu D. The survey on arsenism caused by drinking water. *Chung Hua Liu Hsing Ping Hsueh Tsa Chih* 1993; 14: 195-8.

Zhang L, Chen C. Geographic distribution and exposure population of drinking water with high concentration of arsenic in China. *Wei Sheng Yen Chiu* 1997; 26: 310-3.

PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2001; 10: 149–156
DOI: 10.1002/pds.573

ORIGINAL REPORT

Survey of knowledge and practice on oral contraceptive and emergency contraceptive pills of drugstore personnel in Hat Yai, Thailand

Chaveewan Ratanajamit*[1] MSc(Pharm) and Virasakdi Chongsuvivatwong[2] MD, PhD

[1]*Department of Clinical Pharmacy, Faculty of Pharmaceutical Sciences, Prince of Songkla University, Thailand*
[2]*Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand*

SUMMARY

In Thailand, oral contraceptive (OC) and emergency contraceptive pill (ECP) are available as over-the-counter (OTC) drugs, and drugstores share 30% of services. While the rate of dispensing contraceptive pills has increased, the knowledge and awareness of ECP use is limited among users and providers. The objective of this study was to assess knowledge and practice of drugstore personnel on providing OC and ECP, in order to improve the quality of services. Drugstores located in Hat Yai District, Songkhla Province, Southern Thailand, were the accessible population. There were 109 drugstores, half of them owned by pharmacists. The population was stratified by owner (pharmacist or non-pharmacist) and randomly selected to obtain a sample size of 30 drugstores for each class. Two study methods, questionnaire interview and secret shopping, were used to measure knowledge, and practice, respectively. History-taking, drug-choosing, and advice-giving were the domains measured. The results demonstrated that knowledge on OC was fair, but that on ECP was poor. Pharmacists had better knowledge of proper history taking and ECP indication than non-pharmacists. OC and ECP provision were inappropriately practised in drugstores in the study area. A majority of drugstores were mainly owned by non-pharmacists. For OC practice, drug-choosing was good, but history-taking and advice-giving were poor in both groups. Although both groups dispensed ECP poorly, pharmacists dispensed significantly better than non-pharmacists. Among non-pharmacist staff, the average scores of OC advice-giving, and ECP dispensing, were statistically significantly better among those working in pharmacist-owned drugstores. Both knowledge and practice on OC and ECP should be improved in both types of drugstores in the study area. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS — ECP; OC; drugstore; pharmacist; emergency contraception

INTRODUCTION

Oral contraception (OC) is the most popular fertility regulating method worldwide,[1,2] both in developed and in developing countries.[1,3] The status of OC in most developed countries is prescription-only,[1,4,5] while it is available as an over-the-counter (OTC) drug in most developing countries.[6,7] The emergency contraceptive pill (ECP) is available as a prescription drug in developed countries with intensive campaigns launched to promote its availability and use.[8,9] In most developing countries, there is no policy to promote the use of ECP.[10] In the countries where OC and ECP are available as prescription-only drugs, the need for OC[4,5] and ECP[11–13] availability OTC are increasing. In addition, a pilot project making ECP available OTC is being run in the US.[11,14]

In Thailand, drugstores share approximately 30% of the OC service.[3] The rates of ECP use are not available, but a study on the guidance for the use of

*Correspondence to: Chaveewan Ratanajamit, Department of Clinical Pharmacy, Faculty of Pharmaceutical Sciences, Prince of Songkla University, Songkhla Province, 90110 Thailand. Tel: 66-074-429754. Fax: 66-074-212900.
E-mail: rchaveew@ratree.psu.ac.th

Contract/grant sponsor: Thailand Research Fund.

Received 8 February 2000
Accepted 30 November 2000

levonorgestrel (the available ECP product in Thailand) among drugstore personnel, reported that this drug has been used widely and inappropriately among teenagers (N. Matananpun *et al*., unpublished data). Therefore, it is necessary to conduct a survey of the knowledge of and practice regarding OC and ECP provision by drugstore personnel since they play an important role in providing OC and ECP to teenagers and other users.

The aim of this study was to measure the knowledge and practice of OC and ECP provision by drugstore personnel, to compare the quality of service given by drugstores owned by a pharmacist and by a non-pharmacist and to measure the association between knowledge and practices on OC and ECP services.

STUDY METHODS

Setting

In Thailand, there are approximately 10,000 drugstores. Of these, approximately 45% are registered as being run by pharmacists [15] who may or may not be the owner. In practice, very often drug dispensing is carried out by a non-pharmacist employee.

The study was carried out in 1999 in Hat Yai, the largest city in southern Thailand with a population of 150,000 and 109 drugstores, 55 of which were pharmacist-owned.

Sampling

The drugstores were classified according to whether or not the owner was a pharmacist (for the pharmacist-owned drugstores, the pharmacist is assumed to spend more time with the client and closer supervision of other personnel than a (usually part-time) pharmacist employee in the non-pharmacist owned drugstores). Thirty pharmacist-owned and 30 non-pharmacist-owned drugstores were randomly selected from each class.

Ethical clearance

The study protocol was approved by the ethical committee of the Faculty of Medicine, Prince of Songkla University.

Development of instrument

For OC, a checklist for good practice guidelines was produced by the Ministry of Public Health. This includes the need to interview for contraindications and physical examination including blood pressure measurement, but does not include PAP smear. For ECP, there is no such a guideline. The checklist for ECP was developed based on the fact that it should be used within defined time constraints to prevent pregnancy when no contraceptive had previously been used or when the method failed. The questionnaire for interviewing the dispenser was devised using the checklist. Questions were devised to obtain the answers which reflected each item on the checklist. A weighting scheme was set up by the research team. The questionnaire and the weighting scheme were then discussed with a panel of pharmaceutical experts in the university and further modification to the weighting scheme was made as deemed appropriate.

Data collection

Drugstore visit and interview: Formal visit to and interview of drugstore personnel were carried out by the research team. A pre-tested questionnaire, including closed-ended and open-ended questions, was used to measure the knowledge, regarding both OCs and on ECPs, of volunteer drugstore personnel. It recorded the sociodemographic data of the interviewee, the most common OC dispensed to a new user, whether history taking was routinely carried out (and if so the details such as contraindications, menstrual cycle history, contraceptive use and concurrent drug use), the major concern about the content (type and amount of estrogen and progestin) of the OC chosen, the frequency and content of advice given (such as OC dosing and continuation, common side-effects and problem solving, danger signs of OC). Regarding knowledge about ECP, the questionnaire sought to assess whether sufficient details of the index coitus were taken from the client, the ability to evaluate risk of pregnancy from the unprotected intercourse, the type of ECP dispensed and related advice, and the ECP regimens available in the drugstore for different indications.

Secret shopping. Secret shopping study was conducted after finishing the formal interview by a group of well-trained and validated research assistants (two females and two males) who were pharmacy students. In each drugstore, both OC and ECP were bought by one male shopper and one female shopper on different occasions at least 1 day apart using the scenarios given in Appendix.[1] During shopping, the main variables related to the behaviour of the seller were noticed and

memorized by the shopper. The data were then immediately recorded into the checklist form after the shopper left the drugstore.

Data management and analysis

The data were coded after checking for completeness and accuracy and then computerized using Epi Info version 6.03 and analysed using Stata version 6.0. Data from the questionnaire interview and from secret shopping were handled in the same manner. Variables were grouped into four categories; history taking, drug-choosing and advice-giving for both OC and ECP and another category that was different between OC and ECP (see Table 3). The items on history-taking, drug-choosing and advice-giving were similar in both OC and ECP, while items related to OC choosing, which were of most concern (to examine their knowledge about the differences between high-dose versus low-dose pills, especially in terms of composition, drug of choice for OC initiator and adverse effects) and that on ECP indication were different and not related. The knowledge about what history (e.g. medical, menstrual, concurrent drug use) should be asked, what drug should be chosen and on the basis of what, what and how advice should be given to the new client was asked in the questionnaire interview. In secret shopping, it measured whether these components were actually practiced or not. Each item was scored 1 for a correct answer or practice and 0 for an incorrect answer or practice. The score was then multiplied by a weighting factor (shown in Table 1) whose value was based on its relative importance. For example, in the interview, history taking on disease contraindication was considered to be the most important and was given a weighting of 20, whereas menstrual cycle history was believed to be relatively less important and was weighted at 15, etc. From secret shopping, the weighting scheme was slightly different from the questionnaire. History taking on previous use of OC was considered important because more intensive counselling is needed among new OC users. Disease contraindication had a lower weighting because in our scenario, the client was healthy. Knowledge and practice composite scores were derived from the summation of the weighted score of each item that belonged to the composite. The internal consistency of the items in each category from the questionnaire study were checked using Cronbach's alpha which ranged from 0.47 and 0.94.

Student's *t*-test was used to test for statistically significant differences between groups of sample i.e. owner or interviewee (or seller for secret shopping study). Each composite score was rated on the basis of the average composite score (pharmacist-owned and non-pharmacist-owned drugstore). Multiple linear regression was used to test the independent influence of owner and seller on the score of practice from secret shopping.

History-taking and advice-giving practice in each drugstore in the secret shopping situation were cross-tabulated against corresponding answers in the questionnaire to check the discrepancy between actual and reported practice.

RESULTS

In most pharmacist-owned drugstores, the questionnaires were answered by the owner, but in only half of them did the owners act as sellers in the secret shopping study (Table 2). In non-pharmacist-owned drugstores which were supposed to be run by pharmacists, both questionnaire answering and secret shopping were almost exclusively carried out by non-pharmacists. This indicated that the actual service in a majority of drugstores was mainly provided by non-pharmacists. As expected, informants for the pharmacist-owned drugstore generally had a higher education than those in non-pharmacist-owned drugstores.

From the OC questionnaire study, the only statistically significant difference in composite scores between groups was on advice-giving (pharmacist-owned drugstores did better) (Table 3). The knowledge about OC was acceptable in all groups. The average composite scores for ECP were poor in all groups. Pharmacists had better knowledge of proper history-taking (especially in details of time and frequency of intercourse–not shown in the table) and indications for ECP than the non-pharmacist group. There were no statistically significant differences on ECP choosing and advice giving.

Data from secret shopping of OC and ECP were stratified by both owner and seller (Table 4). Pharmacist sellers in non-pharmacist-owned drugstores had significantly higher scores (mean $\pm$ SD) in OC history-taking (16.2 ± 3.2 vs. 10.0 ± 7.6) and advice-giving (18.0 ± 1.4 vs. 13.7 ± 5.3). There was no problem with OC choosing because most of the available OC in local market were standard low-dose pills at the time of this study. Of several items of quality of service on OC and ECP from secret shopping, only drug dispensing was rated good for OC and fair for ECP. The remaining items were all poor or very poor. Pharmacist sellers generally gave better service than non-pharmacist sellers, except in OC dispensing.

Table 1. Scheme for weighting of items of knowledge and practice

Variable	Questionnaire interview		Secret shopping	
	OC	ECP	OC	ECP
1. History-taking of OC and ECP				
History of use of OC or other methods of contraception	-	-	8	-
The two most important issues on OC history taking		-	-	-
• disease history (contraindication for OC user)	20	-	8	-
• Cycle history	15	-		
• last menstrual period	-	-	3	4
• regularity of cycle	-	-	1.5	1.5
• cycle length	-	-	1.5	1.5
• menstrual duration	-	-	1.5	-
• menstrual related symptoms	-	-	1.5	-
History of concurrent drug use	10	-	6	-
Assessment of hormonal type	5	-		
History of sexual intercourse	-		-	
• Time of intercourse	-	10	-	10
• Time related to menstrual cycle	-	5	-	-
• Frequency of intercourse	-	5	-	-
• Prevention	-	-	-	5
ECP indication	-	-	-	5
2. Drug-choosing				
Popular brand of OC dispensed in drugstore	10	-	-	-
Brand of OC most frequently dispensed to initiator	10	-	-	-
Most concern on pill choosing	10	-	-	-
First OC pill dispensed in secret shopping	-	-	10	-
Actual pills dispensed	-	-	10	-
ECP dispensed	-	-	-	25
Availability of pills within	-		-	
• 1 h of intercourse	-	5	-	-
• 24 h of intercourse	-	5	-	-
• 72 h of intercourse	-	5	-	-
Effectiveness of ECP	-	5	-	-
3. Advice-giving				
Drug dosing for				
• 21/22 tablets package	3.5	-	5	-
• 28 tablets package	3.5	-	5	-
• how to continue the new package	-	-	2	-
• ECP within 1 h of intercourse	-	5	-	-
• ECP within 24 h of intercourse	-	5	-	8
• ECP within 72 h of intercourse	-	5	-	-
Advice on pill missing	5	-	5	-
Side-effects	7	-	7	6
Solving of common side-effects	-	-	4	2
Warning signs for pill stopping	7	-	10	-
Follow advice strictly	-	-	-	2
Evaluation of the effectiveness of ECP	-	-	-	10
ECP indication and limitation	-	-	-	10

Table 5 cross tabulates answers in the questionnaire with data from secret shopping. The strikingly discordant numbers were in ECP history-taking and advice-giving. Among those shops where these practices were claimed in the questionnaire, more than a half did not provide such service during the secret shopping. No statistically significant predictors were found by multiple linear regression.

Table 2. Characteristics of drugstore personnel

Characteristics	Questionnaire		OC shopping		ECP shopping	
	Pharmacist owner $N=30$	Non-pharmacist owner $N=30$	Pharmacist owner $N=60$*	Non-pharmacist owner $N=60$*	Pharmacist owner $N=60$*	Non-pharmacist owner $N=60$*
Provider status						
Pharmacist	26	0	35	2	32	0
Non-pharmacist	4	30	25	58	28	60
Interviewee's education			–	–	–	–
Primary school	0	1				
Secondary school	1	15				
Undergraduate	1	4				
Graduate	22	9				
Postgraduate	6	0				
Average duration of experience (mean ± SD) (years)	5.9 ± 4.0	12.6 ± 10.0	–	–	–	–

*N is twice that in the questionnaire study because each drugstore was visited by one male and one female shopper.

Table 3. Summary of the composite and total scores (mean ± SD) from the OC and ECP questionnaire interview stratified by owner and by interviewee

Variable (maximum score)	Owner		Interviewee		Overall assessment
	Pharmacist ($N=30$)	Non-pharmacist ($N=30$)	Pharmacist ($N=26$)	Non-pharmacist ($N=34$)	
OC					
History taking (20)	15.5 ± 6.1	15.2 ± 5.7	16.3 ± 5.4	14.6 ± 6.2	Good
Most concern on OC choosing (10)	7.0 ± 3.0	5.8 ± 4.4	6.7 ± 3.1	6.2 ± 4.3	Fair
OC choosing (20)	19.2 ± 2.3	18.0 ± 3.8	19.2 ± 2.3	18.1 ± 3.7	Good
Advice giving (26)	17.0 ± 4.5*	13.5 ± 5.6*	17.7 ± 4.3†	13.4 ± 5.4†	Fair
Total score (76)	58.7 ± 9.5*	52.6 ± 13.1*	60.0 ± 8.8*	52.3 ± 11.8*	Fair
ECP					
History taking (20)	11.5 ± 4.9*	8.0 ± 7.0*	12.1 ± 4.9†	7.9 ± 6.6†	Poor
ECP indication (5)	4.5 ± 1.5**	2.0 ± 2.5†	4.8 ± 1.0†	2.1 ± 2.5†	Fair
ECP choosing (20)	11.0 ± 3.3	9.2 ± 2.3	11.2 ± 3.6	9.3 ± 2.2	Poor
Advice giving (15)	4.2 ± 3.7	3.7 ± 2.2	4.2 ± 3.9	3.7 ± 2.2	Poor
Total score (60)	35.7 ± 10.0†	24.8 ± 11.6†	37.1 ± 9.6†	25.0 ± 11.2†	Poor

*$p < 0.05$; †$p < 0.01$, the level of statistically significant difference between different owner and between different interviewee.

DISCUSSION

The results demonstrate that knowledge of OC was good or fair, but knowledge of ECP was poor in all groups. In pharmacist-owned drugstores, very few sellers were pharmacists indicating that these drugstores had only pharmacists' licenses that were required by law. Practice on OC and ECP dispensing were inappropriately undertaken in drugstores in the study area. Both kinds of pills were sold without or with very little history-taking and advice-giving. Failure to practice proper history-taking before OC dispensing might do harm to those who have some contraindications for OC use. Statistically significant higher composite scores on ECP practice such as history-taking, drug-dispensing and advice-giving including OC advice-giving practice in pharmacist-owned drugstores indicated that irrespective of who provided the service, the quality of service was better in pharmacist-owned drugstores. It is very important to counsel clients about pill dosing, common or possible side-effects, pill-by continuation, and drug interaction, otherwise it might result in pill discontinuation, and failure. It has been found that very few

Table 4. Summary of the composite and total scores (mean ± SD) from the OC and ECP secret shopping study stratified by owner and by seller

Variable (maximum score)	Owner						Overall assessment
	Pharmacist seller			Non-pharmacist seller			
	Pharmacist	Non-pharmacist	Total	Pharmacist	Non-pharmacist	Total	
OC	$N=35$	$N=25$	$N=60$	$N=2$	N=58	$N=60$	
History taking (31)	10.0 ± 7.6	8.1 ± 4.8	9.2 ± 6.6	16.2 ± 3.2	7.7 ± 4.5	8.0 ± 4.7	Very poor
Drug dispensing (20)	17.1 ± 6.2	17.2 ± 6.1	17.2 ± 6.1	20.0 ± 0	16.9 ± 6.5	17.0 ± 6.5	Good
Advice giving (32)	13.7 ± 5.3	12.0 ± 4.5	13.0 ± 5.0	18.0 ± 1.4	9.4 ± 4.5	9.7 ± 4.7†	Poor
Total score (83)	40.8 ± 14.5	37.3 ± 9.6	39.3 ± 12.7*	54.2 ± 4.6	34.0 ± 10.4	34.7 ± 10.9*	Very poor
ECP	$N=32$	$N=28$	$N=60$	$N=0$	$N=60$		
History taking (22)	4.3 ± 7.2	2.0 ± 5.4	3.2 ± 6.5*	–	1.2 ± 3.3*		Very poor
Drug dispensing (25)	11.7 ± 10.6	10.7 ± 8.7	11.2 ± 9.7†	–	6.8 ± 6.9†		Poor
Advice giving (28)	7.4 ± 8.3	5.1 ± 5.5	6.3 ± 7.2†	–	3.3 ± 4.7†		Very poor
ECP indication and limitation (10)	6.2 ± 4.9	5.7 ± 5.0	6.0 ± 4.9	–	4.7 ± 5.0		Poor
Total score (85)	29.7 ± 23.6	23.5 ± 14.3	26.8 ± 19.9†	–	16.0 ± 12.4†		Very poor

*$p < 0.05$; †$p < 0.01$, the level of statistically significant difference between different owners.

Table 5. Cross tabulation of answers from questionnaire interview against actual practice in the secret shopping study

Answer in the questionnaire	Actual practice (number of shoppers who received these services)			
	0	1	2	Total
OC				
Advice giving				
Take	1	4	51	56
Not take	0	1	3	4
ECP				
History taking				
Take	37	12	3	52
Not take	8	0	0	8
Advice giving				
Take	33	24	0	57
Not take	3	0	0	3
ECP available (Yuzpe regimen)				
Yes	1	3	0	4
No	52	4	0	56
ECP available (levonorgestrel alone)				
Yes	4	0	0	4
No	40	16	0	56

OC users know the basic pill rules.[16] For ECP practice, history-taking about unprotected intercourse is very important to evaluate the risk of becoming pregnant. ECP counselling (which emphasizes ECP dosage administration, side-effects, and limitations) is still important even though noncompliance may not be a problem.[17]

The knowledge and practice of ECP provision were consistently poor, whereas sellers knew these details of OC well, but practised its administration improperly. Although knowledge and practice in pharmacist-owned drugstores were better than those in non-pharmacist-owned drugstores, both kinds of drugstores should be improved. The knowledge of OC was good, but the selling practice was poor indicating that knowledge was not appropriately applied in practice. It has been documented that the major causes of OC failure (and unwanted pregnancy) were lack of knowledge,[18] noncompliance,[19] and OC discontinuation.[20] These might result from the provider's lack of awareness about the need for counselling. We did not have data on these variables among our clients, but a national contraceptive prevalence survey reported that the 6-month OC discontinuation rate was 27.3%.[3] That might be significantly reduced by proper counselling.

Discrepancies exist between knowledge and actual practice. For example, among the 52 drugstores in which the interviewee answered that history-taking was done before ECP was dispensed, only 15 actually enquired about history-taking in practice and 12 of those did so only in the case of one of the two secret shoppers. In addition, 16 out of 56 drugstores where levonorgestrel-only emergency contraception was reportedly not available (in the questionnaire

interview) sold this drug to our shoppers. In Thailand, the drug company leaflet states that one tablet of levonorgestrel-only emergency contraception should be used within 1 h after unprotected intercourse. This is contrary to the dosing recommendation established by WHO which states that it can be used up to 72 h after unprotected sexual intercourse to prevent pregnancy (either when no method was used or the method failed at intercourse).[21,22] From interview, dispensers' knowledge was based on the leaflet, but they still gave this ECP after 24 h of unprotected intercourse. This discrepancy of knowledge (which was incorrect) and practice (which was against their belief) was also found among obstetrician–gynecologists in a study in Brazil.[23] Answering that this drug was not available in a drugstore did not imply that it would not be sold. High-dose combined contraceptive pills such as the Yuzpe regimen was used as an off-labeling, so that it was not popular among drugstore providers, especially those who were not pharmacists.

The fact that ECP provision practice was poor was not surprising because knowledge was poor. A study conducted by Rojpibulsatit in the same region of Thailand found that knowledge regarding the treatment of common diseases (such as gonorrhoea) was good, but that of uncommon disease (i.e. chancroid) was poor.[24] The poor knowledge and poor practice relating to ECP might be partly explained by this reason. Although knowledge on ECP pharmacology has been available and it has been proved to be safe[25] and effective,[26] it has not reached those who need it throughout the world.[10,27] In addition, it was found that health providers in South Africa were not aware of the dispensing of ECP even though the policy about ECP use in unprotected intercourse was clear.[10]

The higher scores regarding OC advice-giving and ECP dispensing among non-pharmacist sellers working in the pharmacist-owned drugstores than among those working in the non-pharmacist-owned drugstores, reflected the role of pharmacists in training their assistants.

Without adequate counselling on both the advantages and limitations of each type of oral hormonal pill to new clients, unwanted pregnancies will result.

KEY POINTS

(1) Acceptable OC knowledge, but improper practice
(2) Poor ECP knowledge and practice
(3) Pharmacist dispensed ECP better than non-pharmacist

ACKNOWLEDGEMENTS

This study was supported by the Thailand Research Fund. It is a part of the first author's thesis to fulfil the requirements of the PhD degree in Epidemiology at Prince of Songkla University.

REFERENCES

1. Gerstman BB, Burke L, Delaney J, McLellan B. Steroial contraceptive use update, United States, 1989–1994. *Pharmacoepidemiol Drug Safe* 1996; **5**: 141–147.
2. McCormick E. Birth control study: pill is the first choice. *Pharm Times* 1992; **58**: 31–32.
3. Chamratrithirong A, Prasartkul P, Thongthai V, Guest P. National Contraceptive Prevalence Survey 1996. Population and Social Research Institute: Nakorn Pathom, Thailand, 1996; 13–20.
4. Trussell J, Stewart F, Potts M, Guest F, Ellertson C. Should oral contraceptives be available without prescription? *Am J Public Health* 1993; **83**: 1094–1099.
5. Nayak RA, Brushwood DB, Ranelli PL. Consumer views on making oral contraceptives available without a prescription. *APhA Annual Meeting* 1994; **141**: 140.
6. Paxman JM. Roles for non-physicians in fertility regulation: an international overview of legal obstacles and solutions. *Am J Public Health* 1980; **70**: 31–39.
7. Bangboye EA. Oral contraceptive marketing in Ibadan, Nigeria. *Soc Sci Med* 1992; **35**: 903–906.
8. Trussell J, Ball J, Koenig J, Buss M, Allina A, Gamble VN. Call 1-888-NOT-2-LATE: promoting emergency contraception in the United States. *J Am Med Womens Assoc* 1998; **53**: 247–250.
9. Blanchard K. Improving women's access to emergency contraception: innovative information and service delivery strategies. *J Am Med Womens Assoc* 1998; **53**: 238–241.
10. Rutgers RA, Verkuyl DA. Please help, our condom tore last night. *S Afr Med J* 1998; **88**: 143–145.
11. Hutchings J, Winkler JL, Fuller TS *et al*. When the morning after is Sunday: pharmacist prescribing of emergency contraceptive pills. *J Am Med Womens Assoc* 1998; **53**: 30–232.
12. Ellertson C, Trussell J, Stewart FH, Winikoff B. Should emergency contraceptive pills be available without prescription? *J Am Med Womens Assoc* 1998; **53**: 226–9,232.
13. Matheson CI, Smith BH, Flett G *et al*. Over-the-counter emergency contraception: a feasible option. *Fam Pract* 1998; **15**: 38–43.
14. Landis NT. Seattle pilot project makes emergency contraception available directly from pharmacists. *Am J Health Syst Pharm* 1998; **55**: 520–523.
15. Thai FDA. Knowledge about drugs. In: *Handbook for Officers Working for Thai FDA*. Thai FDA: Bangkok, 1997; 262–282.
16. Little P, Griffin S, Kelly J, Dickson N, Sadler C. Effect of educational leaflets and questions on knowledge of contraception in women taking the combined contraceptive pill: randomised controlled trial. *BMJ* 1998; **316**: 1948–1952.

17. Anonymous. ACOG practice patterns. Emergency oral contraception. Number 2, October 1996. American College of Obstetricians and Gynecologists. *Int J Gynaecol Obstet* 1997; **56**: 203–210.
18. Klein J. Emergency contraception. *Hosp Pract Off Ed* 1998; **33**: 26.
19. Davis AJ. The role of hormonal contraception in adolescents. *Am J Obstet Gynecol* 1994; **170**: 1581–1585.
20. Alihonou E, Carre N, Capochichi V, Thonneau P. Contraceptive continuation and its determinants in Benin. *Contraception* 1997; **55**: 97–101.
21. Anonymous. Randomised controlled trial of levonorgestrel versus the Yuzpe regimen of combined oral contraceptives for emergency contraception. Task Force on Postovulatory Methods of Fertility Regulation. *Lancet* 1998; **352**: 428–433.
22. Khanna J (ed). *Improving Methods of Emergency Contraception*. Prog Hum Reprod Res. World Health Organization. Geneva 1999; No. 51: 1–8.
23. Galvao L, Diaz M, Osis MJ, Clark S, Ellertson C. Emergency contraception: klnowledge, attitudes and practices among Brazilian obstetrician-gynecologists. *Int Fam Plan Perspect* 1999; **25**: 168–171.
24. Rojpibulstit M. Drug dispensing for sexually transmitted diseases in drugstores in Songkhla province. Songkla Med J 1999; **16**: 213–22.
25. Glasier A. Safety of emergency contraception. *J Am Med Womens Assoc* 1998; **53**: 219–221.
26. Ho PC, Kwan MS. A prospective randomized comparison of levonorgestrel with the Yuzpe regimen in post-coital contraception. *Hum Reprod* 1993; **8**: 389–392.
27. Gardner J. Emergency contraceptive pills: safe and effective but not widely used. *Wash Pharm* 1997; **39**: 17–21.

APPENDIX 1

Scenario

Oral contraceptive (OC). A recently married woman aged 23 years who plans to have a child in the next 2 or 3 years comes into a drugstore. She has never used any methods of contraception. She is healthy, her medical history is negative for contraindication of combined oral contraceptive(COC) use. Her menstrual history is normal and regular except moderate dysmenorrhea. Is she a good candidate for COC? If yes, what COC should be selected for this client and what instructions/advice should this woman receive regarding use of COC?

Emergency oral contraceptive Pill (ECP). A college student experienced unprotected mid-cycle intercourse, 24 h ago. What is the role of the drugstore in the prevention of an unwanted pregnancy at this time?

IMPLEMENTING THE UNIVERSAL HEALTH COVERAGE: WHICH SOURCE OF INFORMATION IS MORE RELIABLE?

Vinai Leesmidt[1], Supasit Pannarunothai[2] and Virasakdi Chongsuvivatwong[3]

[1]Khlongkhlung Hospital, Khlong Khlung District, Kamphaeng Phet 62120; [2]Center for Health Equity Monitoring, Faculty of Medicine, Naresuan University, Phitsanulok 65000; [3]Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand

Abstract. The implementation of universal health coverage needs accurate data on the distribution of health benefit coverage, particularly the uninsured. The national surveys and routine reports are two important sources of information ready for use. This study shows the validation of data from two sources. The data from national household surveys on the medical welfare, the health card and the social security schemes were validated with the routine report data of the Ministry of Public Health (MOPH) and the Social Security Office (SSO) by provinces. There were considerable differences between these data sets. The national survey data gave a 1.5 times higher estimate than the report data of the MOPH and the SSO. Financial implications of using inaccurate data to implement the universal health coverage could be huge, depending on the capitation rate.

INTRODUCTION

Universal health coverage has become one of the most important health policies in Thailand since the victory of the new government early 2001. The first near-majority victory was the result of a brave election campaign to cover all Thai citizens. The government has to merge the existing scattered, fragmented health coverage scheme in order to cover the uninsured population (see appendix). The most crucial questions for policy implementation were: how many were the uninsured, and how much money the government had to raise in order to achieve universal coverage. These questions need basic information on the coverage of various existing insurance schemes. Unfortunately, the existing information on insurance coverage and its relevant information such as the distribution of the coverage was insufficient, inaccurate and, often, inconsistent.

Basically, information used for health planning can be obtained from either survey or routine report, or both. Particularly, the survey data have become more important if the routine reports are inadequate. This problem is very prevalent in the developing and the least-developed countries (Oyoo *et al*, 1991; Indrayan, 1995). In the same vein, the information about the insurance coverage depends much on the survey data because the data from the report is not complete. However, the survey data are still inconsistent with the report data. So there is a necessity that the data from these two different sources be validated in order to know how much they differ from each other. It is hoped that the result of the data validation will be beneficial for data improvement for future health insurance planning and implementation. This paper aimed to present the methods of data validation and the implications on policy implementation for effective universal health coverage planning and development. These experiences may be learned by other countries attempting the universal coverage policy.

This data validation aimed at identifying the differences between the survey data of the National Statistical Office and the report data of the Ministry of Public Health and the Social Security Office about the insurance coverage under the public medical welfare, health card and social security schemes. Financial implications were estimated if inaccurate data were used for policy implementation.

Correspondence: Dr Supasit Pannarunothai.
Fax: +66 55 261198; E-mail: supasitp@nu.ac.th,

MATERIALS AND METHODS

Sources of data

There were two major sources of data for validation. The first data set was the National Statistical Office survey on household health and welfare 1999. The second data set was the report data to the Ministry of Public Health (MOPH) and the Social Security Office (SSO), Ministry of Labor and Social Welfare.

The survey was the cross-sectional national data obtained from the 32,724 households in May 1999. Sample households were taken from all provinces in the country as the National Statistical Office (NSO) aimed to represent situation of each province. These households accounted for 94,971 individual members (National Statistical Office, 2000).

The routine reports were also the 1999 data on the entitlement to health coverage of the medical welfare and the voluntary health card schemes under the MOPH (Ministry of Public Health, 1999), and the social security scheme under the SSO. The MOPH reports were compiled from the registration system of individual members in 76 provinces including the Bangkok Metropolitan. This data set has been used by the MOPH to allocate the medical welfare and the health card budget according to the number of cards issued to the target populations in each province since 1999. The SSO report was compiled from the database on the insured workers choosing their main contractors at the annual matching process that the insured exercising their choices. The provider choice data have been used to pay main contractors (hospital with 100 beds or higher) on a capitation basis since 1994. Unfortunately, study could not be carried out for the civil servant medical benefit scheme because database for the scheme was non-existent, even though it was the most expensive scheme compared with others (Tangcharoensathien *et al*, 2001).

Data structure and data handling

The unit of analysis of this study was province. Data handling to support provincial analysis had to be accomplished. The format of the survey data differed from those of report data. The structure of the survey data was an individual record of each member in a household (Table 1) while the report data (at MOPH and SSO) were readily available by province (Table 2). Steps were undertaken to make analysis by province possible:

Firstly, all data sources needed to have the same provincial codes in order to link all data to the same province.

Secondly, collapsed the survey data (94,971 individuals) into 76 provincial data according to provincial code. Created new variables to each individual record according to the insurance status. To get the estimates of how many people were under the existing schemes in each province, the sampling fraction (or weight in Table 1) of each individual record was used to blow up to represent the whole population, then sum the blow up by province (Fig 1).

Thirdly, merged data from different sources according to provincial code.Finally, validated the report and the survey data using double log graph (plotting the report data on the X-axis and the survey data on the Y-axis) in STATA (1999). Presented the discrepancy by descriptive statistics such as means and standard deviation.

RESULTS

The results will be presented in 2 parts. First, the same data from two different sources were validated to establish the discrepancy index. Financial implications were then estimated to flag the warning if inaccurate data were used for implementation.

Data validation

The matched data were validated respectively according to the insurance schemes. The first data validation started with the distribution of insurance coverage under the medical welfare scheme. Fig 2 showed that in most provinces the coverage of the medical welfare lied below the line of identity (if both sources of data had the same value, the scatter plots

Table 1
Data structure of the survey data.

Record	Cwd*	Insurance*	Weight*
1	10	4	911.73
2	10	3	1,057.71
3	10	1	993.53
4	10	8	112.01
:	:	:	:
30675	71	6	123.94
30676	71	4	106.73
:	:	:	:
94971	76	5	55.22

* Cwd = Provincial code
* 1 = Civil servant medical benefit scheme
3 = Social security scheme
4 = Medical welfare scheme
5 = Voluntary health card scheme
6 = Private insurance
8 = Uninsured
*The sampling fraction or weight was the figure given by the National Statistical Office according to different sampling proportions in urban, semi-urban and rural areas.
Source: National Statistical Office, 1999.

Table 2
Data structure of the report data.

Record	Cwd	WC	HC	SSS
1	10	780,794	5,427	174,533
2	11	202,435	26,379	636,056
3	12	155,700	97,144	162,991
4	13	116,682	25,242	248,757
:	:	:	:	:
62	77	136,154	21,875	41,431
63	80	637,826	30,597	41,508
:	:	:	:	:
75	95	194,435	18,608	19,160
76	96	328,555	6,158	10,916

Cwd = Provincial code; WC = Welfare card; HC = Health card; SSS = Social security scheme.
Source: Ministry of Public Health, 1999, and the Social Security Office, 1999

would lie along this line). The mean of the medical welfare coverage in all provinces was 2.4 below the identity line. It could be either the routine data were over-reported, or the survey data under-reported. However it indicated that the data from the national survey and the report of the MOPH differed from each other.

Table 3 summarized the differences between the survey and report data by type of health benefit schemes. In contrast to the medical welfare scheme, the survey data gave a 5.4 times higher than the MOPH report data for the voluntary health card scheme. Since respondents of the NSO survey may be confused between the medical welfare and the health card schemes, the two schemes were then combined and analysed on the log graph. By this time, the differences became narrower. The survey under-reported about 1.6 times lower than the MOPH data.

The closest variation between two sources of data was the social security scheme. However, the survey gave the lower estimate, average at 1.5 times lower than the SSO data (Fig 3).

When subcategories of the medical welfare scheme were analysed according to age (0-12, 13-59 and 60 years and above), the elderly gave the narrowest variation (1.4 times) while the working age had the widest variation (10 times). When analysed the variations by region (76 provinces were collapsed to 5 regions, namely: Bangkok, the North, Northeast, Central and South), it was surprising that the disagreement became wider, *eg* the disagreement of the medical welfare rose from 2.4 to 3.2. On the contrary, the disagreement of the social security scheme when analysed by region gave the opposite direction as when analysed by province. The survey data gave higher estimate than the SSO data. However, if looked at the scatter graph in Fig 4, it was more surprising that the respondents in Bangkok produced a higher disagreement (between the survey and the report) than respondents of other regions.

Financial implications

Financial implications of using inaccurate data for implementation of the universal cov-

Rec	Cwd	Ins	Wt
1	10	8	911.73
2	10	4	1,057.71
:	:	:	:
2,000	32	5	1,218.18
:	:	:	:
94,971	96	3	912.85

Create new variables from ins

Rec	Cwd	WC	HC	SSS	Un	Wt
1	10	0	0	0	1	911.73
2	10	1	0	0	0	1,057.71
:	:	:	:	:	:	:
2,000	32	0	1	0	0	1,218.18
:	:	:	:	:	:	:
94,971	96	0	0	1	0	912.85

Blow up data (insurance type x wt)

Rec	Cwd	WC	HC	SSS	Un	Wt
1	10	0	0	0	911.73	911.73
2	10	1,057.71	0	0	0	1,057.71
:	:	:	:	:	:	:
2,000	32	0	1,218.18	0	0	1,218.18
:	:	:	:	:	:	:
94,971	96	0	0	912.85	0	912.85

Data processing

Sum by Cwd

Rec	Cwd	WC	HC	SSS	Un
1	10	103,200	35,835	1,463,910	4,546,725
2	11	26,566	8,594	418,319	648,479
3	12	13,618	2,618	123,420	432,315
:	:	:	:	:	:
62	77	61,645	41,789	7,302	199,869
63	80	431,466	176,318	45,849	725,635
:	:	:	:	:	:
75	95	153,085	129,538	10,898	88,488
76	96	289,377	164,192	6,193	323,506

Fig 1–Handling of data from the National Statistical Office for validation.

erage were illustrated. To achieve universal coverage, the uninsured population was the target to be brought under the umbrella of the universal health insurance. Therefore the degree of financial implication varied according to the number of the uninsured and the capitation rate for health benefit coverage.

The number and percentage coverage of the insurance schemes estimated from the national survey and the routine reports were shown in Table 4. It was assumed that the number of the uninsured was also 1.5 times higher in the survey data than in the reports (since the survey produced lower coverage of benefit schemes, the uninsured therefore tended to be overestimated, 1.5 times was used for the least difference). If 24.8 million uninsured was used to estimate the additional budget to achieve universal coverage, the budgets would vary from 6.8 billion baht to 37.2 billion depending on the capitation rate (Table 5). However, if the number of the uninsured was 1.5 times lower than the survey figure, the range of the additional budget was narrower (from 4.5 billion to 24.8 billion baht).

Table 3
Summary results of the validation of insurance schemes.

	No.	Direction	Mean	SD	Min-Max
By schemes					
Medical welfare	76	Low*	2.37	1.84	0.82-13.32
Health card	76	High	5.40	2.95	0.30-14.10
Medical welfare + health card	76	Low	1.58	1.73	0.27-11.56
Social security	76	Low	1.54	1.17	0.12-6.71
Medical welfare by age group					
0-12	76	Low	3.19	4.86	0.80-30.30
13-59	76	Low	9.95	23.07	0.60-179.0
60+	76	Low	1.36	0.81	0.62-6.43
By region					
Medical welfare	5	Low	3.21	2.87	1.53-8.28
Social security	5	High	2.33	3.43	0.68-8.46

*The survey was lower than the routine report.

Table 4
Insurance coverage distribution 1999.

Type of insurance	Survey		Report	
	Population	Percentage	Population	Percentage
Social security scheme	4,319,083	7.00	4,079,128	6.61
Medical welfare	13,990,560	22.67	23,181,057	37.57
Health card	11,094,440	17.98	2,305,154	3.76
CSMBS	5,485,784	8.89		
Private insurance	834,806	1.35	32,139,342 (CSMBS+Private Insurance+Other +Uninsured)	52.09 (CSMBS+Private Insurance+Other+ Uninsured)
Other	1,057,858	1.71		
Uninsured	24,808,433	40.21		
Blank	113,616	0.18		
Total	61,704,581	100	61,704,581	100

Source: National Statistical Office 1999, Ministry of Public Health 1999, Social Security Office, 1999.

Table 5
Scenarios of financial implication for universal health insurance financing.

Per capitation (Baht)	Survey		Expected*		Discrepancy	
	Uninsured (million)	Budget (million)	Uninsured (million)	Budget (million)	Uninsured (million)	Budget (million)
273[b]	24.8	6,770	16.5	4,505	8.3	2,266
1,197[c]	24.8	29,686	16.5	19,751	8.3	9,935
1,500[d]	24.8	37,200	16.5	24,750	8.3	12,450

* = uninsured of expected numbers is 1.5 lower than the survey data.
[b] = per capitation of medical welfare scheme.
[c] = per capitation agreed by the Bureau of Budget.
[d] = per capitation of the universal coverage proposed by the working group.

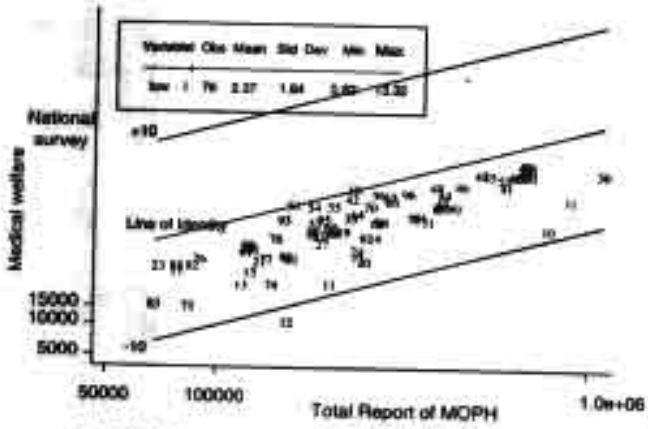


Fig 2–Validation of log of total insurance coverage of the medical welfare scheme.

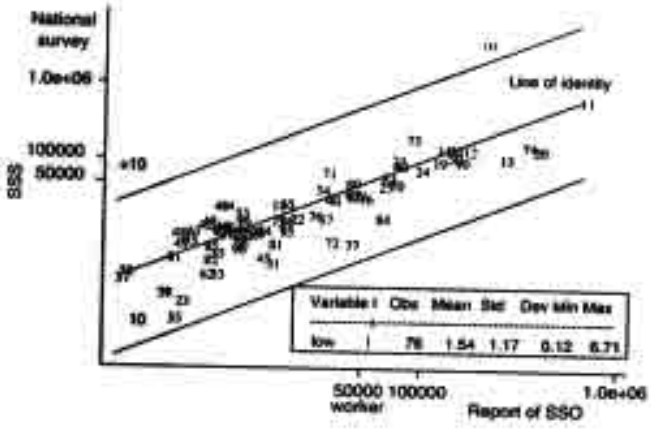


Fig 3–Validation of log of total insurance coverage of the social security scheme.

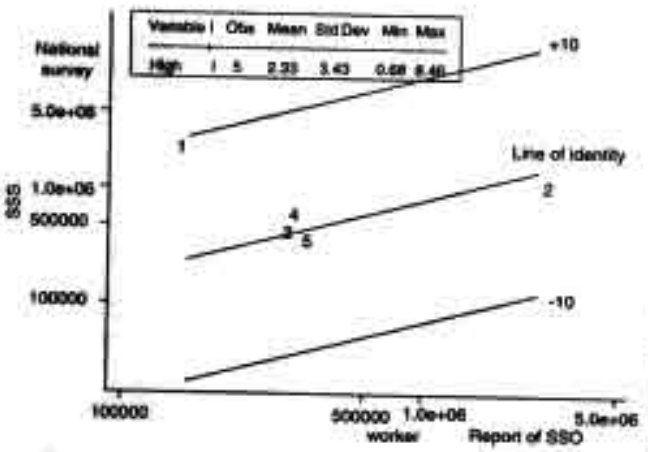


Fig 4–Validation of log of insurance coverage of the social security scheme by region.

Three scenarios of different per capitation rates were used. The lowest capitation rate (273 baht as for the medical welfare scheme in 2001) produced a discrepancy of 2.3 billion baht between the survey and the expected figures. If the capitation rate increased to 1,197 baht as agreed by the Bureau of Budget at the meeting chaired by the Prime Minister on 17 March 2001 (Ministry of Public Health, 2001), the discrepancy would be 9.9 billion baht. If the capitation rate was 1,500 baht as recommended by the Working Group of the Health Systems Research Institute (2001), the additional budget varied from 24.8 billion to 37.2 billion, or with the discrepancy of 12.5 billion baht.

Inaccuracies of the target population affected the financing of both health care providers and the government. If the expected data was used for universal health insurance implementation, it may create financial problem to the services facilities and hospitals, because there would be some degree of under-reporting leading to financial inadequacy. But if the survey data were used, it would be the great burden for the government to raise the subsidy. Therefore, it was important for the government to obtain more accurate data on the population coverage before implementing the universal health insurance policy. Moreover, it was as well important to justify the optimal level for per capitation rate that would bring about the most efficient resource use in order to meet the real health needs.

DISCUSSION

To consider the data validation of the medical welfare scheme in Fig 2, it was found that most of the insurance coverage distribution was below the identity line. It might be either the result of under-report of the national survey or over-report of the MOPH. It was more likely to be under-reported because the respondents who responded to the questionnaires did not know exactly what types of insurance their household members were entitled to. In addition, misclassification was very likely between the medical welfare and the

voluntary health card as both respondents and interviewer had little knowledge on the rapidly changing policy of the MOPH. The reason to support the over-reporting of MOPH system because the numbers of the coverage of the medical welfare were used for budget allocation of the scheme, therefore, most provinces tended to report the maximum number of the target group to maximize the budget allocation. While the process of issuing the welfare card was slow, therefore the reported data were higher and people who had received the welfare cards were lower than the target.

The pattern for the health card coverage was the opposite of the medical welfare scheme. It might be the effect of over-report of the national survey because of misclassification between the health card and the medical welfare card. In other way, it might be the under-report of the MOPH. But the under-report was unlikely because the process of health card issuance was checked by the report of the banking system. The transactions through the bank account were used for allocating budget subsidy. Therefore, the national survey was more likely to be over-reported.

Two reasons to explain why the routine report of social security insurance was very close to the survey were as follows. The routine data were derived from the pay role system. The contributions paid by employees, employers and the government were collected monthly to the social security fund. Furthermore, all social security members had to register with the main contractor for their entitlement, therefore the report of the SSO was more accurate and more reliable than the MOPH's medical and health card schemes.

When the medical welfare distribution was combined with the health card scheme, the pattern of the coverage shifted, and came closer to the identity line. It could be that the effect of misclassification between the health card and the medical welfare card had cancelled each other. The mean of the difference changed from 2.2 lower for the medical welfare and 4.9 higher for the health card to be 1.6 lower the identity line, very close to the social security scheme.

Within the medical welfare insurance, there were several sub-groups from age-related (children and the elderly), income-related (the low income), social characteristics (the veteran, community and religious leaders) and disability-related (the handicapped). If re-categorized the medical welfare into three main groups according to their age (0-12 years, 13-59 years and 60 years and over), children and the elderly were straightforward for their eligibility, therefore the disagreement of both information sources was narrow. The working age group produced the highest variation between the survey and report data. Issuance of the medical welfare card for the working age had to rely on income means testing, which was more difficult than calculating the age.

Comparing the distributions of medical welfare, and social security schemes by region and by province, the medical welfare by region produced a higher discrepancy but with the same directions as those of provincial distribution. However, the social security scheme produced results with different directions when analysed by province and by region. This was because the analysis by region gave equal weight to the 5 regions, while analysis by province gave equal weight for 76 provinces. So the result by region was biased in favor of Bangkok but not for the result by province.

Conclusions and recommendations

From the results of the data validation above, some conclusions could be drawn as follows:

- It is evident that there were considerable differences between the survey data of the National Statistical Office and the report data of the Ministry of Public Health and the Social Security Office.

- The differences varied according to the insurance types of which the medical welfare and voluntary health card had higher variations than the social security insurance.

- The national survey data had a tendency to be over-reported, on average 1.5 times over

the report data.

- Misclassification between medical welfare and voluntary health card led to the variation of the distributions in different age groups and regions of the country.

The followings were recommendations for the implementation of universal health coverage policy:

- It would be dangerous if the universal health coverage policy would be set up and implemented by using single source of data either the survey or the report data.

- It would be crucial that the data from various sources and be validated before finally used.

- It would be necessary to improve the data quality before using them for the implementation of universal coverage policy to reduce errors and weaknesses of the existing data.

- Routine data are a necessity for the proper monitoring system such as financial audit of the policy, however survey data are also helpful for confirmation of the reliability of the data.

Appendix

In 1999, the Health and Welfare survey of the National Statistical Office showed that the civil servant medical benefit scheme which was financed by taxation as a fringe benefit for government employees covered about 9% of the total population. The compulsory social security scheme for employees in formal private sectors covered about 7% of the total population. The MOPH had introduced the low income card scheme to cover the poor, the elderly, children under 12 years old, the handicapped, community and religious leaders about 23% of the total population. The voluntary health card scheme of the MOPH covered family who purchased the health card annually about 18% of the total population. Because 40% of the population were left uncovered, therefore the new government proposed the universal coverage policy. This policy aimed to cover all populations in informal sectors by consolidating previous MOPH's schemes (see Table).

Schemes before 2001	% of total population	From 2001 onwards
Civil servant medical benefit scheme	9	Same target
Social security scheme	7	Same target
Formal sector employees	16	Same
Low income card scheme	23	
Voluntary health card scheme	18	The new universal coverage scheme
Private health insurance	1	
Others	2	
Uncovered	40	
The informal sectors	84	
Grand total (61,704,581 pop)	100	Totally covered

Source: Analysis on the Health and Welfare Survey, 1999.

ACKNOWLEDGEMENTS

This study was partly supported by the Royal Golden Jubilee Program of the Thailand Research Fund.

REFERENCES

Indrayan A. Informatics: the key to efficiency. *World Health Forum* 1995; 16: 305-312.

Ministry of Public Health. Handbook of Guidelines for Health Card Fund. Health Insurance Office. Nonthaburi: Ministry of Public Health, 1999.

Ministry of Public Health. Financial requirements for implementing universal coverage. Paper presented at the meeting on 17 March 2001 chaired by the Prime Minister at the Government Office, Bangkok, 2001.

National Statistical Office. *Report of the Health and Welfare Survey 1999*. Bangkok: The Prime Minister Office, 2000.

Oyoo A, Burstrom B, Forsberg B and Makhulo J. Rapid feedback from household surveys in PHC planning: an example from Kenya. *Health Polic Plann* 1991; 6: 380-3.

Tangcharoensathien V, Srithamrongsawat S, Pitayarangsarit S. Health insurance system – An overview. In: Wibulpolprasert S, ed. Health insurance system in Thailand. Nonthaburi: Health Systems Research Institute, 2001.

STATA Corporation. STATA Release 6: STATA Reference Manual Release 6. Texas: STATA Press College Station, 1999.

The Working Group on Universal Coverage, Health Systems Research Institute. Universal coverage. Nonthaburi: Health Systems Research Institute, 2001.

Original Article

Effects of haemoglobin and serum ferritin on cognitive function in school children

Rassamee Sungthong[1] MD, Ladda Mo-suwan[2] MD and Virasakdi Chongsuvivatwong[1] MD, PhD

[1]*Epidemiology Unit, Faculty of Medicine, Prince of Songkla University Songkhla, Thailand*
[2]*Enterology and Nutrition Division, Department of Pediatrics, Prince of Songkla University, Songkhla, Thailand*

The association between iron deficiency anaemia and cognitive function impairment has been widely reported in young children, but whether the impairment is a result of iron deficiency per se or a combination of iron deficiency and anaemia, and how these conditions interact, is still questionable. Four hundred and twenty-seven school children from two schools in socioeconomically deprived communities were selected in southern Thailand. Iron status was determined by haemoglobin and serum ferritin concentrations. Cognitive function in this study was measured by IQ test and school performance, including Thai language and mathematics scores, using z-scores based on distributions within the same grade and school. Data on demography and socio-economic status were collected by questionnaire answered by the parents. Linear regression models were used to investigate the effect of anaemia and iron deficiency, reflected by haemoglobin and serum ferritin concentration, on cognitive function and school performance. We found that cognitive function increased with increased haemoglobin concentration in children with iron deficiency, but did not change with haemoglobin concentration in children with normal serum ferritin level. Children with iron deficiency anaemia had consistently the poorest cognitive function (IQ, 74.6 points; Thai language score, 0.3 SD below average; and mathematics score, 0.5 SD below average). Children with non-anaemic iron deficiency but with high haemoglobin levels had significantly high cognitive function (IQ, 86.5 points; Thai language score, 0.8 SD above average; and mathematics score, 1.1 SD above average). This study found a dose–response relationship between haemoglobin and cognitive function in children with iron deficiency, whereas no similar evidence was found in iron sufficient children.

Key words: cognitive function, educational achievement, Hat Yai, IQ, iron deficiency, iron status, school performance, Songkhla, Thailand.

Introduction

Iron deficiency and iron deficiency anaemia (IDA) are common in young children. Cognitive function impairment, the consequence of greatest concern, is well established in school children with late stage iron deficiency once anaemia is recognized,[1–5] but it is still controversial in non-anemic iron deficient children.[2,3,6–8] It has been suggested that tissue iron deficiency may develop early by means of a decrease in iron storage without anaemia that may have non-haematological consequences, for example, cognitive function or physical performance impairment.[6,7] However, previous studies have not demonstrated clearly whether the cognitive function impairment found in children with IDA is due to iron deficiency or a combination of iron deficiency and anaemic status, or how these two conditions interact.

The National Anemia Surveillance Program[9] showed that the prevalence of anaemia has been declining in Thailand. A careful re-evaluation of existing iron supplementation in school children is therefore needed. In this study the baseline data of an intervention study were analysed in order to elucidate the relationships between iron deficiency, reflected by serum ferritin (SF), and anaemia, reflected by haemoglobin (Hb), and their effects on cognitive function in school children. We attempted to separate the effect of iron storage from that of haemoglobin level.

Materials and methods

Study site

From the records of primary schools outside Hat Yai municipality, two primary schools were selected as our study site as they were considered to have a high risk of anaemia, comprised at least 150 students, was accessible by automobile and teachers were willing to co-operate in the study.

These two schools, located approximately 35 km from the research centre, had a mixture of Buddhist and Muslim

Correspondence address: Dr Rassamee Sungthong, Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand.
Tel: +66 74 4511657; Fax: +66 74 212900
Email: sburapat@medicine.psu.ac.th
Accepted 16 July 2001

students. Thai-Buddhists speak Thai whereas Thai-Muslims speak both the Thai and Melayu languages. Family incomes were derived mainly from selling plantation products, particularly the rubber latex. The area is free from malaria and no iron supplementation program had been implemented previously at this study site.

Study subjects

An invitation letter providing information about the research project and requesting consent was sent to the parents of children who were in the first to sixth grades. Only children with written informed parental consent were recruited for the study.

Cognitive function assessment

The cognitive function of each subject was measured using the Test of Nonverbal Intelligence (TONI II)[10] and school performance without knowing the child's iron status. TONI II is easy to administer in the field. It is a standardized and culturally fair test for measuring abstract and figural problem solving, a major aspect of intelligence, in 5–80 years-olds.[10,11] Its high test–retest reliability (R^2 = 0.86) and its high correlation (R^2 = 0.7) with a standard intelligence test (i.e., Wechsler's Intelligence Scale for Children (WISC)) in Thai children was its justification for use in this study.[12]

Each of three trained testers established a rapport with each child and demonstrated how to complete the TONI II (form A). Items are arranged in order of difficulty and each consists of one picture question and four or five available responses, but only one correct response. Each child began with the item indicated in the manual (depending on age) and continued until he/she made three incorrect responses in five consecutive items. Total raw scores were computed and then converted to the corresponding IQ using the table provided within the manual. School performance included the average Thai language and mathematics scores collected from the two latest examinations in the previous academic year.

Determination of iron status

Haemoglobin and SF were used to determine iron status from a single venipuncture. A 2 mL sample of blood was placed into an ethylenediaminetetraacetic acid (EDTA) prepared tube for Hb, and 3 mL was kept in a plastic sealed test tube at ambient temperature (approximately 25°C) for serum ferritin measurement. Within three hours, all blood samples were transferred to the laboratory at Songklanagarind Hospital.

Haemoglobin content was measured using an automated machine (Technicon H*1E™ system; Technicon, Tarrytown, NY, USA) using the cyanmethemoglobin method.[13] Serum ferritin concentration (SF) was assessed by the IMx® assay (Abbott Laboratories, Abbot Park, IL, USA) using the Microparticle Enzyme Immunoassay (MEIA) method.[14]

Collection of other independent variables

Demographic variables, including school, class, sex, age, ethnic group, number of siblings and child ordinal position, and socioeconomic variables, including parents' education in years, father's and mother's occupations (none, casual/farmer/trader, or government officer/private), family monthly income (≤ 5000 baht or > 5000 baht) were collected by a questionnaire answered by parents.

Body weight and height of children wearing school uniforms, without belts or shoes and with empty pockets, were measured using a beam balance Detecto scale and stadiometer (Detecto Scales, Brooklyn, NY, USA) to the nearest 0.1 kg and 0.5 cm, respectively. The weight-to-height ratio of each child was then compared with the weight-to-height data from the Nutrition Division, Ministry of Public Health, in 1996–97.[15] Using cut-off points at the 10th, 90th and 97th percentiles for weight-to-height, the children were classified as underweight (< 10th percentile), normal (10th – 90th percentile), overweight (> 90th – 97th percentile) and obese (> 97th percentile). Physical examination was performed by the first author to detect serious illness or infection.

Ethical consideration

This research was approved by the Ethical Review Committee of the Faculty of Medicine, Prince of Songkla University, Thailand.

Data analysis

Scatter plots were constructed between each cognitive function and Hb, broken down by SF into low SF group (≤ 20 µg/L) and normal SF group (SF > 20 µg/L).[16] On the horizontal axis, the points were grouped into three equal bands. The median values of each band were used to fit a cubic spline smoothing function curve.[17]

For modelling, Hb concentration was classified into three groups at the cut-off points of 11.5 and 12.5 g/dL. Mean IQ, Thai language score and mathematics score converted to z-scores, based on the distribution within the same grade and school were compared across each subgroup of Hb and SF, with adjustment for potential confounders. Test for trend was carried out to examine the dose–response relationship between Hb and cognitive function for each group of SF. All analyses were carried out using STATA statistical software version 6 (StataCorp, College Station, TX, USA).[18]

Results

Out of 427 eligible children, the ratio of male to female subjects was 1:1, the average age was 9.6 years and the percentage of underweight and normal weight were 15 and 80%, respectively. Two-thirds of the children were Muslim and slightly more than one half had three siblings or more. Almost all parents were farmers whose formal education averaged 6 years and whose monthly income was less than 5000 baht or US$125.00, compared to the Thai national average of 12 729 baht or $US318.00, reported by the Household Socioeconomic Survey, National Statistical Office. None of the children had overt manifestation of thalassemia disease.

Of the total number of subjects, one-eighth had SF ≤ 20 µg/L and 22% were classified as anaemic (Hb < 11.5 g/dL for 5–11 years old; Hb < 12.0 g/dL for 12–13 years old).[19] The prevalence of IDA in these subjects

was 4.2% of all of the children and 19.4% among the anaemic children. A screening test showed that two-thirds were positive for thalassemia traits.

As seen in Table 1, age, height, parents' education and family monthly income were associated significantly with IQ. Sex, parents' education and family monthly income were associated with Thai language scores whereas only sex and parents' education were associated with mathematics scores. The overall average IQ was 78 ± 12 points (mean ± SD); higher IQ scores were found among children who were older and taller and whose parents had higher education levels and whose families had higher monthly incomes. Better school performance was found among children who were female and whose parents had a higher level education.

Figure 1a,c,e shows that IQ, Thai language and mathematics scores increased with increasing Hb in the low SF group with a significant dose–response relationship, while these cognitive function scores barely changed with Hb concentration in the normal SF group (Fig. 1b,d,f). The group with the highest scores were children with high Hb but low SF. This pattern was verified after adjustment for potential confounders (Table 2).

Using a group of exclusively SF > 20 μg/L subjects as the reference, significantly poorer mathematics score were found in children with low Hb and low SF concentrations, whereas significantly higher of IQ, Thai language and mathematics scores were found in children with high Hb but low SF.

Discussion

There was a significant different pattern of association between cognitive function and Hb in the different iron status groups, as reflected by SF level. IQ, Thai language scores and mathematics scores in children with low SF increased with increasing of Hb with a significant dose–response relationship. In contrast, these measures of cognitive function in children with SF above 20 μg/L was not affected by Hb concentration. Unexpectedly, the highest scores for cognitive

Table 1. Characteristics of children in the study

Variables	n	IQ (point)	Thai language (Z-score)	Mathematics (Z-score)
Categorical variables†				
Sex	427			
Male	199	79 ± 11	−0.4 ± 1.0	−0.2 ± 1.0
Female	228	78 ± 13	0.4 ± 1.0¶	0.2 ± 1.0¶
Weight for height	427			
Underweight	66	78 ± 12	−0.01 ± 1.0	−0.03 ± 1.1
Normal	341	78 ± 12	0.03 ± 1.0	0.02 ± 1.0
Overweight	11	79 ± 8	0.2 ± 1.2	0.24 ± 1.1
Obese	9	77 ± 11	−0.1 ± 0.8	−0.3 ± 0.8
Ethnic group	427			
Thai-Buddhist	133	78 ± 11	−0.01 ± 1.0	−0.1 ± 0.9
Thai-Muslim	294	79 ± 13	0.04 ± 1.0	0.04 ± 1.0
Child ordinal position	426			
≤3	308	78 ± 11	−0.1 ± 1.0	−0.04 ± 1.0
>3	118	79 ± 14	0.2 ± 1.0	0.2 ± 1.0
Siblings	426			
≤3	203	79 ± 12	0.1 ± 1.0	0.1 ± 1.0
>3	223	78 ± 12	−0.02 ± 1.0	−0.1 ± 1.0
Mother's occupation	424			
None	14	81 ± 13	0.3 ± 1.0	0.6 ± 1.0
Casual/farmer/seller	401	78 ± 12	0.01 ± 1.0	−0.01 ± 1.0
Government/officer/private	9	80 ± 10	0.5 ± 1.0	0.6 ± 1.0
Father's occupation	423			
None	4	72 ± 5	−1.0 ± 1.0	0.2 ± 1.0
Casual/farmer/seller	393	78 ± 12	−0.02 ± 1.0	−0.03 ± 1.0
Government/officer/private	26	84 ± 9	0.6 ± 1.0	1.0 ± 1.0
Family monthly income§	418			
≤5000 baht	338	78 ± 12¶	−0.03 ± 1.0¶	−0.03 ± 1.0
>5000 baht	80	82 ± 12	0.2 ± 1.0	0.2 ± 1.0
Continuous variables‡				
Age (years)	427	0.8 ± 0.3¶	−0.1 ± 0.03	−0.05 ± 0.03
Weight (kg)	427	0.2 ± 0.1	−0.004 ± 0.01	−0.01 ± 0.01
Height (cm)	427	0.2 ± 0.1¶	0.001 ± 0.01	−0.001 ± 0.01
Parents' education (years)	418	0.9 ± 0.2¶	0.1 ± 0.2¶	0.1 ± 0.02¶

†Mean ± SD; ‡Coefficient ± SE; §Family monthly income, 1 baht = US$0.025 at the time of data collection; ¶Statistically significant association ($P < 0.05$).

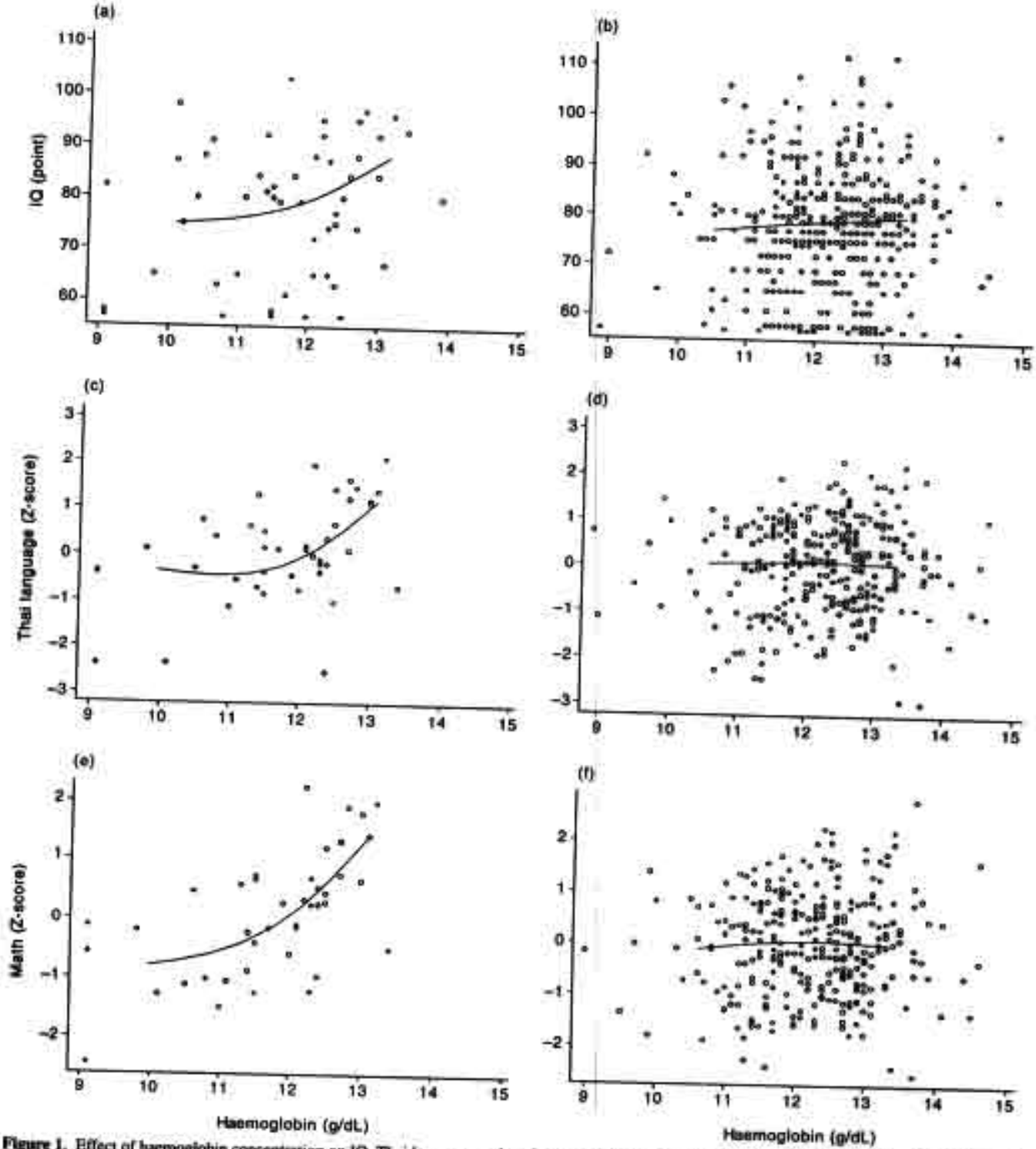


Figure 1. Effect of haemoglobin concentration on IQ, Thai language and mathematics scores in children with (a, c, e) low serum ferritin (≤ 20 μg/L) and (b, d, f) normal serum ferritin (> 20 μg/L) concentrations.

function were in children with high Hb but low SF, whereas the lowest levels were found in those with both low Hb and low SF.

High prevalence of anaemia with relatively low prevalence of iron deficiency (SF ≤ 20 μg/L) suggests that IDA might not be the most common cause of anaemia in this area. Linpisarn *et al.*[20] also found that only 19% of anaemia cases in preschool-age children was IDA, and no IDA was found among anaemic school-age children in Northern Thailand. An SF concentration of ≤ 20 μg/L may not be a gold standard for iron deficiency as it can be elevated with infection or inflammation.[21,22] However, we did not find such infection

Table 2. Mean ± SE of IQ, Thai language and mathematics scores by serum ferritin and haemoglobin categories

Serum ferritin (μg/L)	Hb (g/dL)			P
	<11.5 (n)	11.5–12.5 (n)	>12.5 (n)	
Adjusted mean IQ points†				
≤20	75.0 ± 2.6 (20)	76.2 ± 2.6 (20)	86.5 ± 3.6* (11)	0.03
>20	78.5 ± 1.3 (80)	78.2 ± 1.0 (149)	78.7 ± 1.0 (135)	0.41
Adjusted mean Thai language score (z-score)‡				
≤20	−0.3 ± 0.2 (17)	−0.1 ± 0.2 (15)	0.8 ± 0.3* (9)	<0.01
>20	−0.1 ± 0.1 (64)	0.1 ± 0.1 (119)	0.03 ± 0.1 (125)	0.57
Adjusted mean Mathematics score (z-score)				
≤20	−0.5 ± 0.2* (17)	0.2 ± 0.2 (15)	1.1 ± 0.3* (9)	<0.01
>20	−0.1 ± 0.1 (64)	0.1 ± 0.1 (119)	−0.01 ± 0.1 (125)	0.67

*Significantly different from serum ferritin concentration > 20 μg/L; †Adjusted for family monthly income and parents' education; ‡Adjusted for sex and parents' education. Statistical significance (test for trend): *P*-values < 0.05 were considered statistically significant.

among our subjects. Nopparatana *et al.* found that 24 and 30% of pregnant women and their spouses who attended the antenatal clinic at the teaching hospital near the study area in 1994–95 had thalassemia traits.[23] In addition, β-thalassemia trait, a local common variant, can protect the carrier from iron deficiency.[24] Thus, the thalassemia trait may be an important contributor to non-iron deficiency anaemia. However, due to budget limitations, Hb typing was not undertaken in this study.

Average IQ on figural problem solving, measured by TONI II in this study (mean, 78; SD, 12 point), was relatively low compared with the average IQ in Southern rural Thai children reported by The National Health Survey (mean, 92; SD, 13 point) using similar measurements.[12] This may be influenced by socioeconomic deprivation in the study area.

Although cognitive function tended to be lower in children with iron deficiency anaemia, only mathematics scores were significantly lower than in the normal SF group. Previous studies reported both significantly and non-significantly different lower IQ scores in children with IDA.[2–4,25,26] The variability may be due to using different measurements, in which different aspects of IQ may be assessed. In contrast to our study, Pollitt *et al.*[2] found that Thai language scores were significantly lower while mathematics scores showed no significant difference in IDA children. This inconsistent finding may be explained by differences in assessment systems at different schools. Owing to the use of different measurements of IQ and assessments of school achievement, we may conclude only that IDA has a significant adverse effect at least in some areas of cognitive function.

The adverse effect of iron deficiency on cognitive function may be explained by diminished synthesis, packaging, uptake and degradation of neurotransmitters.[27] The most prominent feature of iron deficiency on cognitive function is the significant and selective diminution of central dopamine neurotransmission.[28] Electroencephalogram power spectrum also showed a slower activity in children with iron deficiency than in iron replete children, suggesting a developmental lag or central nervous system dysfunction among the former group.[29]

Most studies have reported that children with iron depletion did not differ significantly in cognitive function from their peers,[2,3,8,30] whereas a few studies have reported significantly poorer cognitive function.[6,7] Our study unexpectedly found that children with iron depletion accompanying high Hb had the best cognitive function. When Hb is high enough, increasing iron levels may have some adverse effect on cognitive function. Studies in rats and mice reveal that iron overload may cause oxidative stress and has been related to carcinogenesis of the oesophagus and liver.[31–35] In children, iron supplementation does not always result in improvement of cognition.[2,25] This evidence and the findings of our study suggest a need to carefully review the strategy of the iron supplementation program in Thailand.

Bias, chance and confounders can distort the results of epidemiological studies. Measurement bias in this study was overcome by blinding and a standardized process. The investigation of Hb and SF in this study was performed in a standard laboratory using standardized techniques (coefficient of variation = 0.5–5.3%). Low SF has 95–100% specificity in determining low iron stores.[21,36] This study allowed only a 5% chance of type I error, which is acceptably low, and we also found similar associations in different cognitive function tests. It is therefore unlikely to happen by chance. Potential confounders in this study (including socioeconomic status) were adjusted for.

Although bias was minimised and confounders were adjusted for, these results should be interpreted with caution because of the cross-sectional nature of the study. Further randomised control trials should investigate whether an increase in Hb can increase cognitive function in children with iron depletion.

Acknowledgements. We gratefully acknowledge the consultation on haematological issues provided by Assistant Professor Dr Malida Pornpatkul and Dr Malai Wongchaochailert of Songklanagarind Hospital, and the training in the use of TONI II by Assistant Professor Dr Nichara Ruangdaraganon of Ramathibodi Hospital. Special thanks go to Dr Alan F. Geater for his useful comments on the manuscript.

This study forms part of the dissertation of the first author to fulfil the requirement for PhD in Epidemiology at Prince of Songkla University. It is financially supported by the Thailand Research Fund through a Royal Golden Jubilee PhD Program award to Associate Professor Dr Ladda Mo-suwan and Dr Rassamee Sungthong, a Senior Research

Scholar award to Dr Virasakdi Chongsuvivatwong, and a research grant from the Graduate School, Prince of Songkla University.

References

1. Webb TE, Oski FA. Syracuse NY. Iron deficiency anemia and scholastic achievement in young adolescents. J Pediatr 1973; 82: 827–830.
2. Pollitt E, Hathirut P, Kotchabhakdi NJ, Missell L, Valyasevi A. Iron deficiency and educational achievement in Thailand. Am J Clin Nutr 1989; 50: 687–697.
3. Soewondo S, Hussaini M, Pollitt E. Effect of iron deficiency on attention and learning processes in preschool children: Bandung, Indonesia. Am J Clin Nutr 1989; 50: 667–674.
4. Seshadri S, Golpadas T. Impact of iron supplementation on cognitive functions in preschool and school-aged children: the Indian experience. Am J Clin Nutr 1989; 50: 675–686.
5. Soemantri AG. Preliminary findings on iron supplementation and learning achievement of rural Indonesian children. Am J Clin Nutr 1989; 50: 698–702.
6. Bruner AB, Joffe A, Duggan AK, Casella JF, Brandt J. Randomised study of cognitive effects of iron supplementation in non-anaemic iron-deficient adolescent girls. Lancet 1996; 348: 992–996.
7. Zhu YI, Haas JD. Iron depletion without anemia and physical performance in young women. Am J Clin Nutr 1997; 66: 334–341.
8. Walter T, Kovalskys J, Stekel A. Effect of mild iron deficiency on infant mental development scores. J Pediatr 1983; 102: 519–522.
9. Department of Health. Surveillance of anemia in Thailand 1988–97. Bangkok: Ministry of Public Health, 1997; 1–7.
10. Brown L, Sherbenou R, Johnsen S. Test of Nonverbal Intelligence, 2nd edn. Austin, TX: Pro-Ed, 1990; 1–59.
11. Resnick LB, Glaser R. Problem solving and intelligence. In: Resnick LB, ed. The Nature of Intelligence. Hillsdale, NJ: Erlbaum, 1976; 205–230.
12. Chooprapawan J. National Health Survey 1996–97. Bangkok: Health System Research Institute, 1998; 25–57.
13. Technicon. Hemoglobin method. Reference manual Technicon H*1/H*1E system. New York: Miles, 1991; 5,1-11–5,1-13.
14. Abbott Laboratories. IMX System Operation Manual. Abbott Park, IL: Abbott Laboratories, 1996; 1–8.
15. Department of Health. National Reference Value for Weight, Height, and Nutritional Indices for Thai Population Aged 1 day–19 years. Bangkok: Ministry of Public Health, 1999.
16. Winichagoon P, Chongsuvivatwong V. Iron deficiency and iron deficiency anemia. Thailand: Ministry of Public Health, 1991; 1–11.
17. Stata Corporation. Connect-connect points. Stata Graphics Manual Release 6. Texas: Stata Press, 1999; 44–55.
18. Stata Corporation. Stata Statistical Software: Release 6.0. Texas: Stata Press, 1999.
19. Stoltzfus R, Dreyfuss M. Guidelines for the Use of Iron Supplements to Prevent and Treat Iron Deficiency Anemia. Washington DC: ILSI Press, 1998.
20. Linpisarn S, Tienboon P, Promtet N, Putsyainunt P, Santawanpat S, Fuchs GJ. Iron deficiency and anaemia in children with a high prevalence of haemoglobinopathies: implications for screening. Int J Epidemiol 1996; 25: 1262–1266.
21. Mazza J, Barr R, McDonald JWD, Valberg LS. Usefulness of the serum ferritin concentration in the detection of iron deficiency in a general hospital. Can Med Assoc J 1978; 119: 884–886.
22. De-ji T, Ya-min L, Yi L. Serum ferritin in normal subjects and in some diseases. Chinese Med 1987; 100: 631–635.
23. Nopparatana C, Fukumaki Y, Pornpatkul M. Thalassemia in the South of Thailand: Prenatal and Postnatal Diagnosis. Songkhla, Thailand: Department of Pathology, Faculty of Medicine, Prince of Songkla University, 1996; 13–14.
24. Mehta BC. Iron deficiency: prevalence and problems. Part II. J Assoc Physicians India 1990; 38: 491–494.
25. Soemantri AG, Pollitt E, Kim I. Iron deficiency anemia and educational achievement. Am J Clin Nutr 1985; 42: 1221–1228.
26. Pollitt E. Iron deficiency and educational deficiency. Nutr Rev 1997; 55: 133–141.
27. Beard J, Connor JR, Jones BC. Iron in the brain. Nutr Rev 1993; 51: 157–170.
28. Youdim MBH, Schachar DB, Yehuda S. Putative biological mechanisms of the effect of iron deficiency on brain biochemistry and behavior. Am J Clin Nutr 1989; 50: 607–617.
29. Otero GA, Aguirre DM, Porcayo R, Fernandez T. Psychological and electroencephalographic study in school children with iron deficiency. Int J Neurosci 1999; 99: 113–121.
30. Oski FA, Honig AS, Helu B, Howanit ZP. Effects of iron therapy on behavior performance in nonanemic, iron-deficient infants. Pediatrics 1983; 71: 877–880.
31. Chen X, Ding YW, Yang G, Bondoc F, Lee MJ, Yang CS. Oxidative damage in an esophageal adenocarcinoma model with rats. Carcinogenesis 2000; 21: 257–263.
32. Chen X, Yang G, Ding WY, Bondoc F, Curtis SK, Yang CS. An esophagogastroduodenal anastomosis model for esophageal adenocarcinogenesis in rats and enhancement by iron overload. Carcinogenesis 1999; 20: 1801–1808.
33. Pigeon C, Turlin B, Lancu TC, Leruyer P, Le Lan J, Dengnier Y, Brissot P, Loreal O. Carbonyl-iron supplementation induces hepatocyte nuclear changes in BALB/CJ male mice. J Hepatol 1999; 30: 926–934.
34. Abalea V, Cillard J, Dubos MP, Anger JP, Cillard P, Morel I. Iron-induced oxidative DNA damage and its repair in primary rat hepatocyte culture. Carcinogenesis 1998; 19: 1053–1059.
35. Plummer JL, MacKinnon M, Cmielewski PL, Williams P, Ahern MJ, Ilsley AH, de la M Hall P. Dose-related effects of dietary iron supplementation in producing hepatic iron overload in rats. J Gastroenterol Hepatol 1997; 12: 839–842.
36. Kim SK, Cheong WS, Jun YH, Choi JW, Son BK. Red blood cell indices and iron status according to feeding practices in infants and young children. Acta Paediatr 1996; 85: 139–144.

Tropical Medicine and International Health

VOLUME 7 NO 2 PP 133–139 FEBRUARY 2002

Taking a medical history and using a colour scale during clinical examination of pallor improves detection of anaemia

Mahbub-E-Elahi K. Chowdhury[1], Virasakdi Chongsuvivatwong[2], Alan F. Geater[2], Halida H. Akhter[1] and Than Winn[2]

1 *Bangladesh Institute of Research for Promotion of Essential and Reproductive Health and Technologies (BIRPERHT), Dhaka, Bangladesh*

2 *Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat-Yai, Thailand*

Summary We developed a colour tint scale to use as an aid in the clinical assessment of anaemia by measuring conjunctival pallor. The objectives of this study were to evaluate the accuracy and agreement among observers in detecting anaemia in three sequential phases with incremental information using clinical pallor of different anatomical sites, subsequently adding subjects' medical history for physical symptoms and the colour scale. After training in the application of these three sequential assessments, 12 primary health workers were assigned to independently examine 198 anaemic and 254 non-anaemic pregnant women while blind to the true anaemic status. Their assessments in each phase were then compared with the anaemic status based on haemoglobin level, measured using HemoCue, taken as the gold standard, to determine sensitivity and specificity, and agreements among observers in detecting anaemia were calculated. In the three sequential phases of assessment the sensitivities were 73.8, 78.3, 82.9% and specificities 76.0, 84.7 and 90.9%, respectively. In each subsequent step, the improvements in both the sensitivity and specificity were statistically significant [$P(\chi^2_{McNemar}) < 0.01$]. Kappa statistics for agreement among 12 observers for assessing anaemia in the sequential phases were 0.50, 0.71 and 0.82, respectively. The Spearman rank correlation coefficient between haemoglobin level and the colour scale reading was 0.68 ($P < 0.001$). Taking medical history and incorporating a simple colour tint scale with examination of pallor improved the sensitivity, specificity and agreement for detection of anaemia by health workers.

keywords validation, anaemia, colour scale, medical history, pallor, health worker, Bangladesh

correspondence Mahbub-E-Elahi K. Chowdhury, Epidemiology Unit, Faculty of Medicine, Prince of Songkla University (PSU), Hat-Yai 90110, Thailand. E-mail: birperht@citechco.net

Introduction

High prevalence of anaemia during pregnancy (40–60%) is a major problem in many developing countries (WHO 1992). To assist diagnosis of anaemia in rural areas where laboratory facilities are limited, we need a test which is cheap, convenient to use, less dependent on equipment or reagents, reasonably accurate and which requires no invasive procedure. Among simple reliable methods the copper sulphate method (Pistorius *et al.* 1996) and the Lovibond-undiluted method (van Lerberghe *et al.* 1983) have been recommended by several studies for rural settings. However, both depend either on standard solutions or equipment and often are difficult to use. The WHO haemoglobin colour scale, which requires no reagents, has been found to be unreliable in the field setting (van Lerberghe *et al.* 1983). After eliminating the possible sources of error of the haemoglobin scale, Stott and Lewis (1995) developed a new colour scale that proved reliable in laboratory conditions, but in the field had low specificity (47%) at Hb < = 11 g/dl (Van den Broek *et al.* 1999). All of the above methods require either finger blood pricking or venipuncture, which also may be harmful when the supply of instruments and sterilization are not reliable.

Previous studies have shown that the clinical pallor of an individual anatomical site has either low sensitivity or low specificity for diagnosis of anaemia (Nardone *et al.* 1990; Stoltzfus *et al.* 1999). Using the presence of pallor at several sites with an 'or' option could improve the sensitivity; however, the results are still poor (Gjorup *et al.* 1986; Nardone *et al.* 1990). Strobach *et al.* (1988), after measuring the conjunctival pallor with a colour tint scale, found a high correlation (Spearman rank correlation coefficient = 0.84) between the colour scale reading and blood haemoglobin level. However, in that study the observers were not blinded to subjects' haemoglobin level and the authors did not assess how much improvement they could achieve in detecting anaemia after use of the colour scale. These studies have demonstrated that sensitivity for clinical diagnosis of severe anaemia (Hb < 7 g/dl) is high, although traded-off with a high proportion of false positives. However, as the majority of anaemic women have moderate (Hb = 7–10 g/dl) or mild (Hb > 10 & < 11 g/dl) anaemia (WHO 1993), there is a need to improve the sensitivity in these two groups while maintaining an acceptable specificity.

Beside the presence of pallor, many other physical symptoms result from anaemia, among which fatigue, dizziness and palpitation are common. Decreased work capacity and reduced productivity because of anaemia also have been reported by several studies (Scholz *et al.* 1997; Chowdhury *et al.* 2000). To our knowledge no study has explored the usefulness of these physical symptoms in improving the accuracy of diagnosis of anaemia.

Using the simple PC colour technology we developed a new colour tint scale to evaluate conjunctival pallor for clinical assessment of anaemia for use by health workers. We hypothesized that the incorporation of this colour scale and medical history for symptoms related to anaemia with the physical examination of pallor could improve the accuracy and agreement among observers in detecting anaemia. We also estimated the correlation of this colour scale reading with actual blood haemoglobin level and agreement among observers in reading the scale.

Materials and methods

Study design and setting

We evaluated the accuracy and agreement among trained health workers in clinical diagnosis of anaemia in three sequential phases of assessment with incremental information by (1) examining clinical pallor of different anatomical sites, (2) adding medical history of physical symptoms related to anaemia and (3) incorporating the colour scale for assessing conjunctival pallor. During assessment the health workers were blinded to the true haemoglobin level of each subject. Their assessments were subsequently compared with the subject's true anaemia status according to measured haemoglobin level. Prior to this evaluation study, the health workers were trained to clinically assess anaemia over a 4-day training session. The study took place among outpatient pregnant women at two urban health care facilities in Dhaka, Bangladesh, which served the surrounding middle and lower middle class communities, where prevalence of anaemia among pregnant women was 53–60% (Jahan & Hossain 1998). The study protocol was approved by the technical subcommittee and the human subjects subcommittee of the Bangladesh Institute of Research for Promotion of Essential and Reproductive Health & Technologies (BIRPERHT).

Development of colour scale

Our colour scale contained 13 ordinal shades of pink coded 1–13 with an increment of red tint from nearly paper white to deep pink. We made this colour scale by mixing specific proportions of three basic colours – red, green and blue – using the custom colour feature of Microsoft Power Point 97. For shade 01 of the colour scale we mixed 255 units of red with 245 units of each of green and blue. For each of the subsequent shades we kept red fixed at 255 units and subtracted 10 units of each of green and blue from the preceding shade. As the proportion of green and blue decreased the red tinge on the colour scale increased gradually. We printed out the colour scale on white photo paper using an Epson Stylus Photo 700 Inkjet printer at a resolution of 1440 × 720 dpi. We chose this colour scale among eight different scales, with variant shades of red tint, upon agreement among three senior medical students by matching with the natural colour tint of conjunctiva among 20 patients with different haemoglobin levels at the Hematology Department of Songklanagarind Hospital, Thailand.

Gold standard

The measurement of haemoglobin concentration from finger-tip blood using HemoCue (HemoCue AB, Angelholm, Sweden), which employs the modified azide methaemoglobin method of Vanzetti (1966), was taken as gold standard in our study. For calibration of HemoCue we conducted a small pilot study among 20 outpatients at the Hematology Department of Songklanagarind Hospital in Thailand, comparing the results of HemoCue with those of a Coulter Counter. Pearson's correlation coefficient was 0.98. The mean value of difference (Coulter Counter readings – HemoCue readings) was –0.21 g/dl with standard

deviation (SD) of 0.64 g/dl. An experienced laboratory technician collected blood and operated HemoCue.

Phases of clinical assessment

After each of the three phases of clinical assessment each of the observers made her own judgement on the subject's anaemia status in four categories, viz. severe, moderate, mild and non-anaemia. In phase I, each observer independently examined every subject for presence or absence of pallor on nine different anatomical sites, viz. nail-beds, nail-bed blanching, palm, palmar creases, face, lips, tongue, oral mucosa and conjunctiva. In Phase II, in addition to knowledge of clinical pallor, the assessment was based on the subject's medical history which was recorded by a nurse by asking questions of current experience of fatigue, dizziness, palpitation, physical constraints in performing usual daily activities and the burden of workload. The interviewer was blinded from the haemoglobin level and the interview was conducted in a room separately from the observers. The interview results were disclosed to the observers after completion of their assessments in Phase I. In Phase III, the observers assessed the conjunctival pallor by selecting the particular shade of the colour tint that most closely matched the palpable conjunctivae after placing the colour scale adjacent to the everted lower eyelid. The examinations were held in a spacious room at the selected facilities in broad daylight with as much natural light as possible. Examinations were conducted without knowledge of the true haemoglobin level of subjects except during the latter part of the training program. Each observer used a separate observation checklist for every subject.

Selection and training of the observers

We recruited and trained 12 observers, four government and eight non-government organization primary health care providers in Bangladesh. On the first day of a 4-day hands-on training, M-E-E. K. Chowdhury, who was present throughout the study period, familiarized observers with the aetiology and management of anaemia. On the following 3 days each of the observers examined 60 pregnant women from three trimesters covering a range of haemoglobin levels. During training each observer clinically examined every subject in two rounds. In the first round, without having any knowledge of the actual haemoglobin level, the observers independently judged anaemia of one subject through the three sequential phases of assessment. In the second round, the true haemoglobin level of the subject was revealed and the observers were allowed to reexamine the patient through the three sequential phases and reassess their judgement. Through this process the observers could correlate the pallor of different anatomical sites, the physical symptoms and the colour scale with the actual haemoglobin level.

Sample sizes

We estimated sample sizes for anaemic and non-anaemic subjects for each trimester separately using the sample size formula for precision of estimation of proportion. Using the commonly taken cut-off of Hb < 11 g/dl (WHO 1972) for anaemic subjects, we considered the desired sensitivity in first, second and third trimesters as 0.80, 0.85 and 0.80, respectively, and for the non-anaemic the desired specificity as 0.75, 0.80 and 0.75, respectively. Seeking 10% absolute precision with 95% CI for each estimate, the minimum required numbers of anaemic subjects for first, second and third trimesters were 61, 49 and 61 and those for non-anaemics were 72, 61 and 72, respectively.

Subject selection

Pregnant women attending each selected facility were recruited consecutively within each trimester. Gestational age was calculated from the date of last menstrual period (LMP) recorded during monthly visits. The exclusion criteria were: unable to recall the LMP date, unwilling to provide fingertip blood, had inflammation in eyes, had jaundice diagnosed symptomatically by yellow eyes or attending the facility with haemorrhage. A nurse screened the subjects for eligibility criteria and asked the patients to participate. Only women who provided free and informed written consent were finally recruited. Every subject was examined independently by each of the observers for clinical detection of anaemia through the three sequential phases of assessment.

Statistical analysis

In analysis we regrouped the clinical assessment of anaemia status by the observers into two levels, anaemic and non-anaemic, where the anaemic group was formed by combining three subgroups of anaemia, viz. severe, moderate and mild. As each subject was examined by 12 observers, we employed the generalized estimating equations approach to logistic regression, which takes into account the repeated measures, to estimate sensitivity, specificity and 95% CI. For agreement among 12 observers in detecting anaemia at each phase of clinical assessment we used kappa for multiple observers (Fleiss 1981). For testing the hypothesis that incremental introduction of methods improved the sensitivity and specificity we divided the dataset into anaemic and non-anaemic groups according to actual haemoglobin level and in each group analysed the discordant pairs between two

sequential phases of assessment using McNemar's χ^2 with 1 d.f. for paired sample. To compare the improvement in sensitivity in subsequent phases of assessment between mildly and moderately anaemic subgroups, we used the Z-test to compare the McNemar χ^2-values. We used the Spearman rank correlation coefficient (Conover 1980) to summarize the relationship between actual haemoglobin level of the subjects and the colour scale reading. For agreement in reading the colour scale we used weighted kappa between pairs of observers. Weighted kappa was used in order to accommodate different degrees of disagreement of reading the ordinal colour scale by the observers. We applied a weight of 1 for perfect agreement, 0.5 for disagreement of one unit and 0 otherwise. Stata Software 6.0 (Statacorp, 1999, Stata Statistical Software: release 6.0, College Station, Texas, USA) was used in analysis.

Results

In the first, second and third trimester we included 36, 101, 61 anaemic and 90, 86, and 78 non-anaemic subjects, respectively. The smaller than planned number of anaemic subjects in the first trimester theoretically reduced our planned precision from 10% to 13% in that group. We excluded five subjects for an eye inflammation, three diagnosed for jaundice and two for unwillingness to have their finger pricked. Twelve observers made a total of 5165 independent assessments for anaemia among 452 subjects. The maximum number of subjects examined by an observer was 452 and the minimum was 343.

Characteristics of the subjects

The mean age of all the subjects was 23.2 years. The mean haemoglobin level of 198 anaemic subjects was 10.0 g/dl (range = 7.7–10.9 g/dl) and that of 254 non-anaemic subjects was 12.1 g/dl (range = 11.0–15.0 g/dl). Of the anaemic subjects 45.4% were moderately and 54.6% were mildly anaemic.

Improvements in sensitivity, specificity and agreements

Sensitivities, specificities and agreements improved consistently over the three sequential phases of clinical assessment based on pallor, medical history and the colour scale in each category of anaemia as seen in Table 1. Details of tabulation for statistical test of significance are shown in Table 2, where the numbers in the cells are the numbers of matched pair observations.

At each phase of clinical assessment the observers could detect a higher proportion of anaemia among the moderate anaemic subgroup than among the mild anaemic subgroup (Table 1). Comparison of these improvements in sensitivity in subsequent phases of assessment between moderately and mildly anaemic subgroups revealed no significant difference in the improvements from phase I to II, but a significantly greater improvement from phase II to III in mildly anaemic subjects [$P(z) < 0.05$].

Association of anaemia status with physical symptoms

Among the anaemic subjects fatigue, dizziness and palpitation were reported by 90.4, 31.8 and 28.8%, respectively, as compared with 52.4, 13.0 and 4.7% among the non-anaemic. Each of these symptoms was significantly associated with anaemic status [$P(\chi^2_{1df}) < 0.001$]. Fifty per cent of the anaemic women said they had constraints in daily activities compared with less than one in 10 (9.3%) among non-anaemic women; this relationship remained

	(I) Clinical pallor	(II) I + Medical history	(III) II + Colour scale
Sensitivity percentage (95% CI)			
Anaemia Hb < 11 g/dl	73.8 (71.4–76.1)	78.3 (75.8–80.6)	82.9 (81.3–84.4)
Moderate anaemia Hb = 7–10 g/dl	83.3 (80.7–85.6)	90.8 (88.9–92.4)	93.0 (91.2–94.4)
Mild anaemia Hb > 10 & < 11 g/dl	62.8 (59.6–66.0)	67.7 (64.3–70.9)	74.6 (72.1–76.9)
Specificity percentage (95% CI)			
Non-anaemic Hb > = 11 g/gl	76.0 (73.9–78.0)	84.7 (82.8–86.4)	90.9 (89.8–91.9)
Kappa	0.50	0.71	0.82

Table 1 Sensitivity, specificity with 95% CI and agreement (kappa) for detecting anaemia at three sequential phases of assessment using clinical pallor, medical history and the colour scale

Table 2 Matched pair analysis results for test of significance for improvement in diagnosing true anaemics (sensitivity) and true non-anaemics (specificity), from (a) Phase I to Phase II and (b) Phase II to Phase III

a. Improvement from Phase I* to Phase II†

	True anaemics		True non-anaemics	
	Dx in Phase II		Dx in Phase II	
Dx in Phase I	Anaemic	Non-anaemic	Anaemic	Non-anaemic
Anaemic	1466	125	310	392
Non-anaemic	293	367	140	2072
P (McNemar χ^2)	< 0.01		< 0.01	

b. Improvement from Phase II† to Phase III‡

	True anaemics		True non-anaemics	
	Dx in Phase III		Dx in Phase III	
Dx in Phase II	Anaemic	Non-anaemic	Anaemic	Non-anaemic
Anaemic	1701	58	207	243
Non-anaemic	166	326	58	2406
P (McNemar χ^2)	< 0.01		< 0.01	

* Dx based on clinical pallor; †Dx based on clinical pallor and medical history; ‡Dx based on clinical pallor, medical history and color scale

significant [$P(\chi^2_{1df}) < 0.001$] after controlling for burden of daily household activities.

Correlation between Hb level and colour scale reading

The Spearman rank correlation coefficient between haemoglobin level and the colour scale reading for conjunctival pallor was 0.68 ($P < 0.001$). For haemoglobin levels < 11 g/dl most assessors had colour scale readings of 9 or below and for Hb > = 11 g/dl most read 10 or above (Table 3).

Agreement among observers for reading the colour scale

Weighted kappa between pairs of observers for agreement in reading the ordinal colour scale ranged between 0.46 and 0.83. For 52 of 66 possible pairs among 12 observers this kappa value lay between 0.61 and 0.80.

Discussion

In this study, augmenting clinical assessment of pallor successively with medical history and the colour scale improved both the sensitivity and specificity in detecting anaemia among pregnant women. The agreement among observers in assessing anaemia also improved successively. Both the correlation of the colour scale measurements with the actual haemoglobin concentration and the agreement among observers for reading the scale were moderately high.

Examining pallor at nine anatomical sites for detection of anaemia, our observers achieved a sensitivity of 74% and a specificity of 76% among pregnant women, which are higher than those reported by Nardone *et al.* (1990) of 65% sensitivity and 71% specificity for the presence of pallor at any one of conjunctiva, face or palm among older hospitalized whites. The improved accuracy in our study

Table 3 Percentage distribution of ordinal colour scale reading among pregnant women with various levels of haemoglobin

	Ordinal colour scale reading						
Hb (g/dl)	*n*	≤ 6	7	8	9	10	≥ 11
≤ 9.9	884	28.1	42.1	16.0	7.0	5.0	1.9
10.0–10.9	1366	0.8	13.3	36.0	25.6	22.2	2.1
11.0–11.9	1444	0.9	1.0	5.8	19.5	62.3	10.5
12.0–12.9	906	0.8	1.6	4.0	12.0	55.7	25.9
≥ 13	562	–	–	4.3	10.1	52.7	32.9

may be a result of the inclusion of more anatomical sites and an overall judgement of anaemia status based on examination of different sites.

Combining the presence of physical symptoms from medical history with the results of examination of pallor improved the sensitivity to 78% and specificity to 85%. In general, symptoms of morbidities reported by patients during history taking are as valuable as those noted from examinations of professionals (DeGowin 1994). Incorporating medical history also augmented agreement among observers in detecting anaemia. Using the printed colour scale in assessing conjunctival pallor appeared to add an objective element in clinical detection of anaemia that helped further improve both sensitivity and specificity to 83 and 91%, respectively. The colour scale was also useful in raising agreement among observers. An important point to note is that in each subsequent phase of clinical assessment relatively greater improvement occurred in specificity than in sensitivity. This reflects the improved ability of each successive phase to correctly diagnose non-anaemic subjects who were inaccurately diagnosed in the preceding phase of assessment.

In our study, at each phase of clinical assessment the observers could better detect moderate anaemia than mild anaemia, which confirms other study findings (Strobach *et al.* 1988; Stoltzfus *et al.* 1999) that clinical examination is more sensitive in identifying severe or moderate anaemia than mild anaemia. However, we could also demonstrate that incorporating the colour scale gave a relatively greater improvement in identifying mild anaemia than moderate anaemia compared with assessment based on pallor and medical history together. This finding may have a public health implication of using our colour scale for its ability in detecting anaemia at early stages when timely treatment may prevent further reduction of haemoglobin level.

The production of our colour tint scale is simple, inexpensive and, after lamination, convenient to use in the field. This colour scale can be reproduced locally at a health facility with a personal computer and a photo-quality colour printer and thus is less dependent on a central supply. One of the difficulties that may arise in the development of this colour scale is that there may be a variation in printed output depending on the type of printer and its colour technology. Further studies should be carried out to investigate the extent of such differences and the effect of different colour scales on sensitivity and specificity. For the above reasons, we recommend that every time upon arrival of a new batch of colour scales, training of the health workers should take place for 'calibration' of the new scale in the community population. Standardized mass production may simplify the application of the colour scale, even though this would mean some dependence on a central supply. Because of the associated problem of fading of the printed colours in the course of time, especially when used in the field, the user should carefully protect the scale from direct sunlight. Moreover, regular replacement of scales may be required.

We selected anaemic and non-anaemic women in all three trimesters and thus covered a full span of gestational age for validation of clinical diagnosis of anaemia. The proportion of excluded subjects was low (about 2%) and they were distributed across all three trimesters. We included a large number of observers for clinical assessment of every subject whereby the likelihood of observer bias was reduced. Examinations performed by paramedics who work in the field is also an advantage in that these results may be applicable to the setting of Bangladesh. On the other hand, in our study the observers performed the test having fresh knowledge immediately after training, which might have improved the results. In real situations the accuracy of clinical assessment by the paramedics may decrease over time, especially in the absence of regular examination of patients; thus, occasional training may be necessary to maintain their level of skills.

The implication of this study is that physical symptoms are useful in improving the accuracy of diagnosis of anaemia and should be included in the training program of the health workers. The colour tint can also improve the accuracy although training in its use will be required. Assimilating this training with the refresher-training course of the health workers may help implement our method of clinical diagnosis of anaemia along with the existing health care program. In countries where resources are limited, this simple technique may provide an effective means of diagnosis of anaemia particularly suited to use in the field.

Acknowledgements

This paper is a part of first author's thesis to fulfil the requirements for his PhD in Epidemiology at Prince of Songkla University (PSU), Thailand, under BIRPERHT/Ford Foundation funding. This article was developed at the Epidemiology Unit, PSU, which is partially supported by Thailand Research Fund. We would like to thank the Program Manager of Radda Barnen and Project Director of Bangladesh Association for Prevention of Septic Abortion at Mirpur in Dhaka, for cooperation in examining subjects in the respective facilities.

References

Chowdhury ME, Akhter HH & Chongsuvivatwong V (2000) Community-based self-reported symptoms of antepartum

morbidities: the health burden and care-seeking patterns of rural Bangladeshi women. *Southeast Asian Journal of Tropical Medicine and Public Health* 31, 598–605.

Conover WJ (1980) *Practical Nonparametric Statistics* 2nd edn. John Wiley & Sons, New York, pp. 213–343.

DeGowin RL (1994) *Degowin and Degowin's Diagnostic Examination* 6th edn., McGraw-Hill, Inc., New York, pp. 14–34.

Fleiss JL (1981) *Statistical Methods for Rates and Proportions* 2nd edn., John Wiley & Sons, New York, pp. 211–236.

Gjorup T, Bugge PM, Hendriksen C & Jensen AM (1986) A critical evaluation of the clinical diagnosis of anaemia. *American Journal of Epidemiology* 124, 657–665.

Jahan K & Hossain F (1998) *Nature and extent of malnutrition in Bangladesh: Bangladesh National Nutrition Survey 1995–96.* Institute of Nutrition and Food Science, University of Dhaka, Dhaka, p. 91.

van Lerberghe W, Keegels G, Cornelis G, Ancona C, Mangelschots E & van Balen H (1983) Haemoglobin measurement: the reliability of some simple techniques for use in a primary health care setting. *Bulletin of the of the World Health Organization* 61, 957–965.

Nardone DA, Roth KM, Mazur DJ & McAfee JH (1990) Usefulness of physical examination in detecting the presence or absence of anaemia. *Archives of Internal Medicine* 150, 201–204.

Pistorius LR, Funk M, Pattinson RC & Howarth GR (1996) Screening for anaemia in pregnancy with copper sulfate densitometry. *International Journal of Gynaecology and Obstetrics* 52, 33–36.

Scholz BD, Gross R, Schultink W & Sastroamidjojo S (1997) Anaemia is associated with reduced productivity of women workers even in less-physically-strenuous tasks. *British Journal of Nutrition* 77, 47–57.

Stoltzfus RJ, Edward-Raj A, Dreyfuss ML *et al.* (1999) Clinical pallor is useful to detect severe anaemia in populations where anaemia is prevalent and severe. *Journal of Nutrition* 129, 1675–1681.

Stott GJ & Lewis SM (1995) A simple and reliable method for estimating haemoglobin. *Bulletin of the of the World Health Organization* 73, 369–373.

Strobach RS, Anderson SK, Doll DC & Ringenberg QS (1988) The value of the physical examination in the diagnosis of anaemia. Correlation of the physical findings and the haemoglobin concentration. *Archives of Internal Medicine* 148, 831–832.

Van den Broek NR, Ntonya C, Mhango E & White SA (1999) Diagnosis of anaemia in pregnancy in rural clinics: assessing the potential of haemoglobin colour scale. *Bulletin of the of the World Health Organization* 77, 15–21.

Vanzetti G (1966) An azide-methaemoglobin method for haemoglobin determination in blood. *Journal of Laboratory and Clinical Medicine* 67, 116–126.

WHO (1972) *Nutritional Anaemias.* Report of a WHO Group of Experts World Health Organization. Technical Report Series 503. World Health Organization, Geneva.

WHO (1992) *The Prevalence of Anaemia in Women: a Tabulation of Available Information.* World Health Organization, Geneva.

WHO (1993) *Prevention and Management of Severe Anaemia in Pregnancy.* World Health Organization, Geneva.

Tropical Medicine and International Health

VOLUME 7 NO 3 PP 271–279 MARCH 2002

Has directly observed treatment improved outcomes for patients with tuberculosis in southern Thailand?

Petchawan Pungrassami[1], Søren P. Johnsen[2], Virasakdi Chongsuvivatwong[3] and Jørn Olsen[4]

1 *Zonal Tuberculosis Centre, Yala, Thailand*
2 *Department of Clinical Epidemiology, Aarhus University Hospital and Aalborg Hospital, Aarhus, Denmark*
3 *Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hatyai, Songkhla, Thailand*
4 *The Danish Epidemiology Science Centre, University of Aarhus, Aarhus, Denmark*

Summary

OBJECTIVE To validate the practice of directly observed treatment (DOT) and evaluate its effect on treatment outcomes.

METHODS This follow-up study conducted in 24 districts in southern Thailand included 411 new, smear-positive, pulmonary tuberculosis (TB) patients who started treatment between February and September 1999. Patients and/or their observers were interviewed about their actual DOT practice during the first 2 months of treatment. Treatment outcomes were evaluated at the end of the second month and at the end of treatment.

RESULTS Of 411 patients, 379 were assigned to DOT but only 68 practised strict DOT for every dose during the first 2 months. Adjusted odds ratios (ORs) for 'no sputum conversion' and 'unsuccessful treatment' were 1.1 (95% CI 0.6–2.1) and 1.3 (95% CI 0.6–2.8), respectively, for those who practised strict DOT *vs.* the rest.

CONCLUSIONS Actual practice of DOT was quite different from what was intended at the assignment. Practice of strict DOT during the first 2 months was not associated with sputum conversion or treatment success in this study area.

keywords tuberculosis, directly observed treatment, compliance, treatment outcome, Thailand

correspondence Petchawan Pungrassami, Zonal Tuberculosis Centre 12, Amphur Muang, Yala 95000, Thailand. E-mail: ppetch@cscoms.com

Introduction

Thailand is among the 23 countries which host 80% of the estimated incident cases of tuberculosis (TB) in the world (World Health Organization 2001a), and Thailand is among the 10 countries with the highest prevalence of primary multidrug resistance (The WHO/IUATLD Global Project 1997). The World Health Organization (WHO) reviewed Thailand's National TB Programme in 1995 and found low cure rates of 17–68% (Ministry of Public Health 1995). WHO (1999) thus recommended a Directly Observed Treatment, Short-course (DOTS) strategy, which involves the following five elements: (1) government commitment, (2) case detection by sputum smear microscopy, (3) standardized short course regimen with DOT for at least the initial 2 months, (4) uninterrupted supply of qualified drug, and (5) standardized recording and reporting system. As an element of DOTS, DOT implies that an observer watches the patient swallow the medicine with the aim of improving patient compliance, which is considered the most serious problem in TB control (Addington 1979).

The National TB Programme has adopted the WHO proposals and planned to cover all 810 districts by the year 2001 (Payanandana *et al.* 1999). Since the implementation of DOTS in 1996, cure rates have apparently improved, but they are still below the WHO target of 85% (Ministry of Public Health 1999).

In spite of worldwide implementation, the efficacy of DOT remains questionable (Volmink *et al.* 2000). Randomized controlled trials have given opposite results, and compliance to the DOT principle has not been reported in these studies (Zwarenstein *et al.* 1998; Kamolratanakul *et al.* 1999; Zwarenstein *et al.* 2000; Walley *et al.* 2001). It is expected to be difficult to maintain the initial allocation

of DOT over several months, and analysis according to the 'intention-to-treat' principle may leave a too limited exposure contrast to produce meaningful results.

Because the use of DOT observers takes away resources from other tasks, it is important to know if DOT works. The aim of this study is to quantify DOT in practice and to estimate the effect of actual DOT practice on treatment outcomes.

Methods

Study design and population

We selected a group of patients currently diagnosed with TB and followed them to document their actual practice of DOT and treatment outcomes. The study area covered a population of about 1.2 million in 24 districts in southern Thailand.

Through the TB registers at the 22 governmental TB clinics (one zonal TB centre, one regional hospital, three general hospitals and 17 community hospitals), we identified all 455 patients who were compatible with WHO definition of new, smear-positive, pulmonary TB (Maher *et al.* 1997), and who started treatment between 1 February and 30 September 1999.

Treatment

All the patients started the treatment with either daily or intermittent short-course drug regimens. The daily regimen combined a 2-month initial phase of four drugs (isoniazid, rifampicin, pyrazinamide and ethambutol) and a 4-month continuation phase of two drugs (isoniazid and rifampicin). A three-times per week treatment with the same drug combination was used only for 21 patients who agreed to take the medicine at the zonal TB centre. According to national guidelines, the treatment was evaluated by sputum examinations at the end of the second, fifth and sixth months. If sputum still contained acid fast bacilli (AFB) at the end of the second month, the initial phase was extended for 1 or 2 more month(s). For patients who could not tolerate the standard treatment, the drug regimen was changed to a 9-month regimen without pyrazinamide or to an 18-month regimen without rifampicin and pyrazinamide.

Initial assignment of DOT observer

The TB clinic staff chose a DOT observer with the patient's agreement and recorded the information on the patient's treatment card. The initially assigned observers were divided into three groups (Ministry of Public Health 1998): health personnel (including staff members of TB clinics, hospital wards and health centres); community members (including village health volunteers, community leaders and friends); and family members (including both close and distant relatives).

Sources of information

Data on exposure, outcomes and potential confounders were obtained from interviews using structured questionnaires and from medical records. The questionnaires were pilot-tested twice before starting the data collection and participants were interviewed at least once and up to four times during the treatment, and/or after the end of the treatment. The interviewers were 24 health professionals working at the TB clinics involved in the study and the first author (P.P.), who selected and trained all interviewers and supervised the interview process. We informed the patients and/or their observers about the aim of our study, confirmed their verbal consent to be interviewed, and emphasized the importance of their answers before interviewing them during the clinic attendance or the ad hoc home visit. Two cases were interviewed solely by telephone because of misappointment and patient migration.

Exposure variables

Two types of DOT were defined, DOT by the initial assignment (assigned DOT) and DOT in actual practice (actual DOT). The patients initially assigned to have any types of observer were grouped as 'assigned DOT', and those without observer assigned as 'assigned no DOT'.

The patients and/or their observers were asked whether the observers actually watched the patients swallow the medicine during the first 2 months (strict DOT). The less strict practices, such as just staying with the patient during drug intake without watching, preparing medicine for patients, and reminding patients about the drug intake were considered as 'not strict DOT'. Three cut-off points were used to define the actual practice of strict DOT: (1) every dose *vs.* some doses or never, (2) more than 50% of the expected doses *vs.* 50% or less, and (3) every dose or some doses *vs.* never. We defined the former group in each cut-off point as 'actual DOT' and the latter group as 'actual no DOT'. All definitions were analysed, however, we only present results according to the first cut-off point.

Outcome variables

We measured the effect of DOT using the following two endpoints: sputum negative conversion at the end of the

second month, and treatment success at the end of the treatment. The WHO definitions of treatment outcomes (Maher *et al.* 1997) were applied to both endpoints. Accordingly, we classified treatment outcomes into six mutually exclusive categories: (1) AFB-negative sputum at time of evaluation; (2) no AFB result at time of evaluation; (3) AFB-positive sputum at time of evaluation; (4) death (died of any cause before time of evaluation); (5) treatment interruption for at least two consecutive months; and (6) transferred out to another TB clinic with unknown sputum result.

At the end of the second month, patients who had AFB-negative sputum were classified as 'sputum conversion', the rest as 'no sputum conversion'. Similarly, patients who received the full course of treatment and had AFB-negative sputum or had no sputum result at the end of the treatment were classified as 'treatment success', the rest as 'unsuccessful treatment'.

Potential confounders

Several potential confounders were considered and divided into three groups: (1) demographic and socio-economic status (gender, age, marital status, ethnic group, formal education, understanding Thai language, occupation, income, feasibility to be free from work or study, independence in travel and number of living places); (2) health services provided (type of TB clinic, drug regimen, and use of fixed dose combination); and (3) disease condition (initial weight, initial AFB result, initial drug resistance, adverse drug effect, HIV/AIDS status, and associated diseases/conditions including diabetes mellitus, cardiovascular disease, cerebrovascular disease, liver cirrhosis, psychosis, alcoholic consumption, drug abuse and imprisonment). The patients without information on income or susceptibility tests were treated as separate groups.

Variable selection and statistical analysis

Univariate analyses were performed by Pearson's chi-square test. Only covariates with at least marginal association with the outcome (sputum conversion *vs.* no sputum conversion and successful treatment *vs.* unsuccessful treatment) (P-value < 0.2) were selected to be tested in the models, as described in the following paragraph.

Four logistic regression models with increasing numbers of covariates were applied to determine the association between the exposure and the outcome (1) without covariates, (2) with inclusion of the first group of covariates, (3) with inclusion of the first and second groups of covariates, and (4) with inclusion of all three groups of covariates. For each step of adding the group of covariates, only covariates with the following criteria were retained in the model: (1) having significant association with the outcome ($P < 0.05$) or (2) having marginal association with the outcome ($P < 0.1$) plus leading to a change of more than 15% of OR for any DOT comparison in the larger model, if removed. Once a variable was included, it was also included in the following model to ensure comparability of the log likelihood.

The confounders for the effect of each type of DOT on each outcome were controlled by multivariate analysis. The results were presented as odds ratios (OR) with 95% confidence intervals (CIs). We performed the likelihood ratio test to check the significance of each covariate group in the models (Kleinbaum 1994), and used the Hosmer–Lemeshow statistic to check the fit of the logistic regression models (Hosmer & Lemeshow 2000). A P-value of less than 0.05 was considered statistically significant in interpreting the resulting models. All analyses were carried out using STATA (StataCorp. 1999).

Results

Of the 455 patients who started treatment, 44 were excluded because the interviewers were unable to establish contact with them or their DOT observers. Compared with the remaining patients, the excluded patients were younger (median age 31 *vs.* 42 years), were more often HIV positive or suffered from AIDS (27 *vs.* 11%), and had poorer treatment outcomes (sputum conversion rate 57 *vs.* 78%; cure rate 30 *vs.* 75%).

The remaining 411 patients were from 6 to 86 years of age (mean 44 years, 95% CI 42–46); and 75% were male. Of the 323 patients who provided information on income, 76% earned less than the official 'minimal daily wage' in the study area (3–3.5 US$). Among the 104 tested, 2% (95% CI 1–3%) had multidrug resistance (resistance to at least isoniazid and rifampicin).

Initial DOT assignment

Of the 411 patients, 379 (92%) had been initially assigned to any type of observer and the remaining 32 patients were not assigned to any, because they refused to accept one (18), because no suitable observers were available (9) or both (5). DOT was more often assigned to patients who were female, who had a living partner, who were treated at the zonal TB centre, who used fixed dose combination, or who had initial drug resistance (Table 1).

Actual DOT practice

Overall, only 68 of 411 patients (17%) practised strict DOT in every dose during the first 2 months, 97 of 393 patients with available information on the following DOT duration practised for more than 50% of expected doses (25%), whereas 154 of 411 (37%) patients took the medicine without being watched by observers. Of 379 patients assigned to DOT, 65 (17%) practised strict DOT for every dose during the first 2 months, 93 of 362 patients with available information on the following DOT duration practised for more than 50% of expected doses (26%), whereas 133 of 379 patients (35%) never practised strict DOT. Practice of strict DOT for every dose during the first 2 months was found more often among the patients who had lower income, who could be free from work/study, had any drug resistance, who were HIV positive had AIDS or other comorbidity, who travelled with others, who did not use fixed-dose combination, or who were treated at the zonal TB centre (Table 1).

Treatment outcomes

Overall sputum conversion and treatment success rates were 78 and 85%, respectively. Sputum conversion rates were lower among male patients and those who were single, divorced or widowed. Treatment success rates were lower among patients who were male, Bhuddist, never travelled alone, were treated at hospitals, did not use

Table 1 Initial DOT assignment and actual DOT practice during the first 2 months by patient characteristics

Patient characteristics	*n*	Assignment at the start DOT/No DOT	Actual practice 2 months DOT/No DOT
Gender			
Male	308	280/28	53/255
Female	103	99/4	15/88
Living partner			
Not having	140	125/15	23/117
Having	271	254/17	45/226
Minimal daily income			
> Minimal daily wage	79	73/6	6/73
< Minimal daily wage	244	222/22	48/196
No information	88	84/4	14/74
Free from work/study			
Not free	307	282/25	42/265
Free	104	97/7	26/78
Dependence in travel			
Need/with other	139	131/8	42/97
Travel alone	272	248/24	26/246
TB clinic			
Community hospital	236	223/13	32/204
General/regional hospital	119	100/19	20/99
Zonal TB centre	56	56/0	16/40
Use of fixed dose combination			
No use	244	221/23	47/197
Use	167	158/9	21/146
HIV/AIDS status			
No	364	334/30	53/311
Yes	47	45/2	15/32
Co-morbidity*			
No	317	294/23	47/270
Yes	94	85/9	21/73
Initial drug resistance			
Sensitive to all four drugs	85	82/3	9/76
Any resistance	19	19/0	6/13
No test	307	278/29	53/254
All patients	411	379/32	68/343

*Co-morbidity: diabetes mellitus, cardiovascular disease, cerebrovascular accident, psychosis, liver cirrhosis, alcoholic, drug abusers, prisoners.

fixed-dose combination, with HIV/AIDS or with other comorbidity (data not shown).

Table 2 shows the proportions of all possible outcomes at the end of the second month and at the end of the treatment by type of DOT. Short- and long-term outcomes were different between 'actual DOT' and 'actual no DOT' groups, not between 'assigned DOT' and 'assigned no DOT' groups.

Multivariate analysis results

No significant differences in risk of no sputum conversion were seen between DOT and No DOT groups, regardless of DOT types or statistical models (Tables 3 and 4). Males had approximately twice the odds of no sputum conversion as female patients. Risks of no sputum conversion were reduced to about half among the patients who had no living partner or who travelled independently (Model 2).

Without the presence of covariates, DOT was positively associated with the chance of treatment success when data were analysed according to the assignment (Table 5), but DOT had opposite effects when analysed according to the actual DOT practice (Table 6). However, no association was found between DOT, either assigned or actual, and chance of treatment success after adjustment for potential confounders (Model 4 in Tables 5 and 6).

Men had an approximately four- to fivefold increased risk of unsuccessful treatment compared with women (Model 4). The risk of unsuccessful treatment was approximately eight times higher among those with HIV/

Table 2 Treatment outcomes at the end of the second month and at the end of the treatment. Results are given according to the assigned DOT at the start of treatment and the actual DOT practice during the first 2 months of the treatment

Time of evaluation	Type of DOT	Group	*n*	Outcomes (row percentage in parenthesis) AFB negative	No AFB result	AFb positive	Death	Default
At the end of the 2nd month	Assigned	No DOT	32	25 (78.1)	3 (9.4)	1 (3.1)	1 (3.1)	2 (6.3)
		DOT	379	296 (78.1)	29 (7.7)	39 (10.3)	10 (2.6)	5 (1.3)
	Actual*	No DOT	343	271 (79.0)	29 (8.5)	34 (9.9)	3 (0.9)	6 (1.8)
		DOT	68	50 (73.5)	3 (4.4)	6 (8.8)	8 (11.8)	1 (1.5)
	Total		411	321 (78.1)	32 (7.8)	40 (9.7)	11 (2.7)	7 (1.7)
At the end of the treatment	Assigned	No DOT	32	19 (59.4)	4 (12.5)	0	4 (12.5)	5 (15.6)
		DOT	379	290 (76.5)	35 (9.2)	5 (1.3)	30 (7.9)	19 (5.0)
	Actual†	No DOT	343	267 (77.8)	30 (8.8)	4 (1.2)	20 (5.8)	22 (6.4)
		DOT	68	42 (61.8)	9 (13.2)	1 (1.5)	14 (20.6)	2 (2.9)
	Total		411	309 (75.2)	39 (9.5)	5 (1.2)	34 (8.3)	24 (5.8)

*$\chi^2_{4\ d.f.} = 26.65$, $P = 0.000$. †$\chi^2_{4\ d.f.} = 19.31$, $P = 0.001$.

Table 3 Odds ratios of 'no sputum conversion' at the end of the second month for 'assigned DOT' vs. 'assigned no DOT' groups after adding various confounder groups

Included variable	Compared group	OR (95% CI) of no sputum conversion† Model 1	Model 2	Model 3
Assigned DOT	DOT/no DOT	1.00 (0.42–2.40)	0.98 (0.40–2.42)	1.08 (0.43–2.71)
Gender	Male/female	–	**2.36 (1.24–4.47)**	**2.44 (1.28–4.65)**
Living partner	No/having	–	**0.54 (0.31–0.92)**	**0.58 (0.33–0.99)**
Independence in travel	Yes/no	–	**0.55 (0.34–0.91)**	**0.54 (0.33–0.90)**
TB clinic*	CH/GH-RH	–	–	1.02 (0.59–1.75)
	ZTC/GH-RH	–	–	0.48 (0.19–1.22)
Log likelihood		−216.03	−208.00	−206.34
Degrees of freedom		1	4	6

*CH = community hospital, GH = general hospital, RH = regional hospital, ZTC = zonal TB centre.
†Model 1 = crude analysis, Model 2 = Model 1 + demographic and socio-economic covariates, Model 3 = Model 2 + covariates related to health services, Model 4 = Model 3 + covariates related to disease condition. Bold indicates $P < 0.05$.

Table 4 Odds ratios of 'no sputum conversion' at the end of the second month for 'actual DOT' *vs.* 'actual no DOT' groups after adding various confounder groups

		OR (95% CI) of no sputum conversion†		
Included variable	Compared group	Model 1	Model 2	Model 3
Actual DOT	DOT/no DOT	1.36 (0.75–2.46)	1.10 (0.58–2.07)	1.20 (0.63–2.29)
Gender	Male/female	–	**2.34 (1.23–4.44)**	**2.40 (1.26–4.57)**
Living partner	No/available	–	**0.54 (0.31–0.92)**	**0.57 (0.33–0.99)**
Independence in travel	Yes/no	–	**0.57 (0.34–0.95)**	**0.57 (0.33–0.96)**
TB clinic	CH/GH-RH	–	–	1.03 (0.60–1.77)
	ZTC/GH-RH	–	–	**0.48 (0.19–1.20)**
Log likelihood		−215.55	−207.96	−206.21
Degrees of freedom		1	4	6

*CH = community hospital, GH = general hospital, RH = regional hospital, ZTC = zonal TB centre.
†Model 1 = crude analysis, Model 2 = Model 1 + demographic and socio-economic covariates, Model 3 = Model 2 + covariates related to health services, Model 4 = Model 3 + covariates related to disease condition. Bold indicates $P < 0.05$.

AIDS, and approximately twofold for those with other comorbidity. By contrast, the risk of unsuccessful treatment was reduced among patients who travelled alone (OR = 0.2), and patients who were treated at a community hospital (OR = 0.4) or the zonal TB centre (OR = 0.1).

Discussion

We did not find a statistically significant effect of DOT on sputum conversion or on treatment success; regardless of whether DOT was analysed according to the initial assignment or according to what was reported by the patients and/or their observers. A lower chance of sputum conversion was seen among male patients, patients who had a living partner, or who never travelled alone. Furthermore, male patients, who never travelled alone, who were treated at a general/regional hospital, and patients with HIV/AIDS, or other comorbidity, had worse chances of treatment success.

The Centers for Disease Control and Prevention in the United States (US Department of Health & Human Services 1994) and WHO 1994, 1995, 1997, 1999) have recommended DOT for all TB patients because it is difficult to predict whether a patient will follow the treatment. They have done so without requesting DOT to be evaluated in practice including measurement of the adherence to the DOT principle. Our findings show that actual DOT practice could be very different from the intended assignment. The result also calls for caution when interpreting the results of the four randomized controlled trials analysed according to the 'intention-to-treat' principle and without information on actual DOT practice

Table 5 Odds ratios of 'unsuccessful treatment' for 'assigned- DOT' *vs.* 'assigned no DOT' groups after adding various confounder groups

		OR (95% CI) of unsuccessful treatment†			
Included variable	Compared group	Model 1	Model 2	Model 3	Model 4
Assigned DOT	DOT/no DOT	**0.42 (0.19–0.97)**	**0.37 (0.15–0.89)**	0.52 (0.21–1.32)	0.45 (0.17–1.17)
Gender	Male/Female	–	**4.90 (1.97–12.16)**	**5.57 (2.21–14.08)**	**4.33 (1.62–11.56)**
Ethnic group	Muslim/Buddhist	–	**0.44 (0.25–0.79)**	**0.40 (0.22–0.73)**	0.54 (0.28–1.05)
Independence in travel	Yes/no	–	**0.23 (0.13–0.42)**	**0.20 (0.11–0.37)**	**0.21 (0.11–0.40)**
TB clinic*	CH/GH-RH	–	–	**0.46 (0.24–0.88)**	**0.44 (0.22–0.88)**
	ZTC/GH-RH	–	–	**0.15 (0.04–0.56)**	**0.10 (0.02–0.40)**
HIV/AIDS	Yes/no	–	–	–	**8.11 (3.74–17.60)**
Other co-morbidity	Yes/no	–	–	–	**2.40 (1.22–4.70)**
Log likelihood		−174.19	−154.21	−148.27	−131.40
Degrees of freedom		1	4	6	8

*CH = community hospital, GH = general hospital, RH = regional hospital, ZTC = zonal TB centre.
†Model 1 = crude analysis, Model 2 = Model 1 + demographic and socio-economic covariates, Model 3 = Model 2 + covariates related to health services, Model 4 = Model 3 + covariates related to disease condition. Bold indicates $P < 0.05$.

Table 6 Odds ratios of 'unsuccessful treatment' for 'actual DOT' *vs.* 'actual no DOT' groups after adding various confounder groups

		OR (95% CI) of unsuccessful treatment†			
Included variable	Compared group	Model 1	Model 2	Model 3	Model 4
Actual DOT	DOT/no DOT	**2.15 (1.15–4.04)**	**1.39 (0.68–2.82)**	1.61 (0.77–3.35)	1.28 (0.58–2.83)
Gender	Male/female	–	**5.09 (2.05–12.66)**	**5.76 (2.28–14.58)**	**4.62 (1.73–12.33)**
Ethnic group	Muslim/Buddhist	–	**0.47 (0.27–0.84)**	**0.40 (0.22–0.74)**	**0.55 (0.29–1.06)**
Independence in travel	Yes/no	–	**0.26 (0.14–0.49)**	**0.23 (0.12–0.43)**	**0.22 (0.11–0.45)**
TB clinic*	CH/GH-RH	–	–	**0.43 (0.23–0.81)**	**0.40 (0.20–0.79)**
	ZTC/GH-RH	–	–	**0.12 (0.03–0.45)**	**0.08 (0.02–0.32)**
HIV/AIDS	Yes/no	–	–	–	**7.52 (3.46–16.34)**
Other comorbidity	Yes/no	–	–	–	**2.44 (1.25–4.77)**
Log likelihood		−173.42	−156.03	−148.39	−132.50
Degrees of freedom		1	4	6	8

*CH = community hospital, GH = general hospital, RH = regional hospital, ZTC = zonal TB centre.
†Model 1 = crude analysis, Model 2 = Model 1 + demographic and socio-economic covariates, Model 3 = Model 2 + covariates related to health services, Model 4 = Model 3 + covariates related to disease condition. Bold indicates $P < 0.05$.

(Zwarenstein *et al.* 1998; Kamolratanakul *et al.* 1999; Zwarenstein *et al.* 2000; Walley *et al.* 2001). The effect measures of these studies are likely biased by a difference between initial assignment and actual practice.

Balasubramanian *et al.* (2000) were the first to report whether patients travelled to the clinic and received the medicine, as recorded by health workers. They found that 27% of 200 patients did not receive DOT and outcomes among this group were significantly worse than those who received DOT. However, they did not provide information on whether actual DOT was practised among those who came. Accordingly, the proportion of not receiving DOT in this study may not be readily comparable with the proportion of no DOT practice among those assigned to DOT in our study (35%). As indicated in our study, many risk factors confounded the association between DOT and treatment outcomes. Hence, the trend toward worse outcomes among those who did not receive DOT or were not assigned to DOT should not be considered reliable evidence of DOT benefit.

Regardless of the cut-off point used to classify the actual practice of DOT (data not shown), the non-significant association of DOT and outcomes and the independent predictors remained the same. This finding challenges not only the policy of universal DOT (Bailey & Sbarbaro 1988; Anna 1993; Salomon *et al.* 1997; Weis 1997; Bayer *et al.* 1998; Heymann *et al.* 1998), but also the recommended period of applying DOT (World Health Organization 1999). Although the results may reflect selection bias or uncontrolled confounding, the results do not rule out that the practice of DOT during the first 2 months may be counter-productive for some and insufficient for others, particularly for the long-term outcome.

On the other hand, the failure to find significant effects of DOT supports the experience-based conclusion that 'DOT is not panacea' (Fujiwara *et al.* 1997) but DOT may be a part of good case management to support TB patients to achieve cure (World Health Organisation 2001b).

By using observational studies in settings where the DOTS programme has been implemented, we faced an expected lack of comparability between the DOT and No-DOT groups. We used logistic regression models with different number of covariate groups in attempt to determine the confounding effect of each group and finally to adjust for the lack of comparability at baseline. Without randomization, we cannot be sure that the DOT and No-DOT groups are comparable, even after adjustment for the included confounders. The study was also analysed using the confounder score approach as described by Miettinen (1976) in order to make the control of potential confounding more efficient and to examine results as a function of baseline risks. The results were virtually unchanged and the effect measure were rather similar in the different risk groups (data not shown). Randomized trials will not solve the problem of compliance to the DOT principle. Practical problems and the patients' rights to choose the treatment modality they prefer will probably limit the scientific value of a randomized trial that aims at quantifying the effect of DOT. During the course of treatment, the patients may gain insight into their own ability to manage the treatment and may move out of the DOT group that would bias the comparison in disfavour of the DOT group. One should therefore not disregard the DOT principle on the basis of our findings but take the results

as a reason for caution. Any DOT services should try to identify the patients who actually will benefit from the DOT principle and the patients who can manage the drug intake on their own.

Although several steps were taken to examine DOT practice, the study may still be subject to exposure misclassification. DOT practice could be over-stated in the interviews because patients and observers saw interviewers as part of the health personnel. We also miss DOT information on 44 cases who more often interrupted the treatment (20 *vs.* 6%) and if they mostly were No-DOT group, we may underestimate the beneficial effect of DOT.

In conclusion, we found no significant improvement in the prognosis for those who practised DOT as part of their treatment for pulmonary TB. Tuberculosis control programmes should probably not focus on DOT without strengthening other strategies of good patient management.

Acknowledgements

This study was financially supported by Thailand Research Fund through the Senior Research Scholar Grant for Dr V. Chongsuvivatwong and the Royal Golden Jubilee PhD Project, and the Danish Epidemiology Science Centre. The activities of the Danish Epidemiology Science Centre are financed by a grant from the Danish National Research Foundation. Transportation facility during the data collection was provided by Zonal TB Centre 12, Regional Office 12 of Communicable Disease Control, Department of Communicable Disease Control, Ministry of Public Health, Thailand. The study is part of the first author's thesis to fulfil the requirements of the PhD degree in epidemiology at Prince of Songkla University, Thailand.

References

Addington WW (1979) Patient compliance: the most serious remaining problem in the control of tuberculosis in the United States. *Chest* 76 (Suppl.), 5741–5743.

Annas GJ (1993) Control of tuberculosis – the law and the public's health. *New England Journal of Medicine* **328**, 585–588.

Bailey WC & Sbarbaro JA (1988) Controversies in pulmonary medicine. All patients should receive directly observed therapy in tuberculosis. *American Review of Respiratory Disease* **138**, 1075–1076.

Balasubramanian VN, Oommen K & Samuel R (2000) DOT or not? Direct observation of anti-tuberculosis treatment and patient outcomes, Kerala State, India. *International Journal of Tuberculosis and Lung Disease* **4**, 409–413.

Bayer R, Stayton C, Desvarieux M *et al.* (1998) Directly observed therapy and treatment completion for tuberculosis in the United States: is universal supervised therapy necessary? *American Journal of Public Health* **88**, 1052–1058.

Fujiwara PI, Larkin C & Frieden TR (1997) Directly observed therapy in New York City. History, implementation, results, and challenges. *Clinics in Chest Medicine* **18**, 135–148.

Heymann SJ, Sell R & Brewer TF (1998) The influence of program acceptability on the effectiveness of public health policy: a study of directly observed therapy for tuberculosis. *American Journal of Public Health* **88**, 442–445.

Hosmer DW Jr & Lemeshow S (2000) Assessing the Fit of the Model. In: *Applied Logistic Regression* 2nd edn (eds NAC Cressie, NI Fisher, IM Johnstone, P Kamolratanakul, H Sawert, S Lerr *et al.*) John Wiley & Sons, New York, pp. 143–202.

Kamolratanakul P, Sawert H, Lertmaharit S *et al.* (1999) Randomized controlled trial of directly observed treatment (DOT) for patients with pulmonary tuberculosis in Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **93**, 552–557.

Kleinbaum DG (1994) Maximum Likelihood Techniques: An Overview. In: *Logistic Regression. A Self-Learning Text.* (Series eds K Dietz, M Gail, K Krickeberg & B Singer) Springer Verlag, New York, pp. 101–124.

Maher D, Chaulet P, Spinaci S & Harries A (1997) *Treatment of Tuberculosis. Guidelines for National Programmes Global Tuberculosis Programme*, World Health Organization, WHO/TB/97.220 2nd edn. Stabilimento Tipografico Ferrero s.r.l., Romanno Canavese (TO).

Miettinen OS (1976) Stratification by a multivariate confounder score. *American Journal of Epidemiology* **104**, 609–620.

Ministry of Public Health Thailand (1998) *Guidelines for National Tuberculosis Control.* Ministry of Public Health, Bangkok.

Ministry of Public Health Thailand & World Health Organization (1995) *The External Review of the National Tuberculosis Programme, Thailand 18–29 June 1995.* WHO, Geneva.

Ministry of Public Health Thailand & World Health Organization (1999) *Second Review of the National Tuberculosis Programme in Thailand.* WHO/CDS/TB/99.273. WHO, Geneva.

Payanandana V, Kladphuang B, Somsong W & Jittimanee S (1999) *Battle Against TB. National Tuberculosis Programme, Thailand, 1999.* Tuberculosis Division, Department of Communicable Disease Control, Ministry of Public Health, Bangkok.

Salomon N, Perlman DC, Rubenstein A *et al.* (1997) Implementation of universal directly observed therapy at a New York City hospital and evaluation of an out-patient directly observed therapy program. *International Journal of Tuberculosis and Lung Disease* **1**, 397–404.

StataCorp (1999) *Stata Statistical Software*: Release 6.0. Stata Corporation, College Station.

The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance 1994–1997 (1997) *Anti-Tuberculosis Drug Resistance in the World.* WHO/TB/97.229. WHO, Geneva.

US Department of Health and Human Services (1994) *Improving Patient Adherence to Tuberculosis Treatment.* Public Health Service, Centers for Disease Control and Prevention, National Center for Prevention Services, Division of Tuberculosis Elimination, Georgia.

Volmink J, Matchaba P & Garner P (2000) Directly observed therapy and treatment adherence. *Lancet* 355, 1345–1350.

Walley JD, Khan AM, Newell JN & Khan MH (2001) Effectiveness of the direct observation component of DOTS for tuberculosis: a randomised controlled trial in Pakistan. *Lancet* 357, 664–669.

Weis SE (1997) Universal directly observed therapy. A treatment strategy for tuberculosis. *Clinics in Chest Medicine* 18, 155–163.

World Health Organization (1994) *TB A Global Emergency: New Attention to an Old Disease*. WHO Report on the TB Epidemic 1994. WHO/TB/94.177. WHO, Geneva.

World Health Organization (1995) *Stop TB at the Source*. WHO Report on the Tuberculosis Epidemic, 1995. WHO/TB/95.183. WHO, Geneva.

World Health Organization (1997) *Use DOT More Widely*. WHO Report on the Tuberculosis Epidemic 1997. WHO/TB/97.224. Koninklijke van Poll Drukkerijen bv, Roosendaal.

World Health Organization (1999) *What Is DOTS? A Guide to Understanding the WHO-recommended TB Control Strategy Known as DOTS*. WHO/CDS/CPC/TB/99.270. WHO, Geneva.

World Health Organization (2001a) *Global Tuberculosis Control*. WHO Report 2001. WHO/CDS/TB/2001.287. WHO, Geneva.

World Health Organization (2001b) *A Revised Framework for Effective Tuberculosis Control*. http://www.stoptb.org/material/RevisedFramework.htm.

Zwarenstein M, Schoeman JH, Vundule C, Lombard CJ & Tatley M (1998) Randomised controlled trial of self-supervised and directly observed treatment of tuberculosis. *Lancet* 352, 1340–1343.

Zwarenstein M, Schoeman JH, Vundule C, Lombard CJ & Tatley M (2000) A randomised controlled trial of lay health workers as direct observer for treatment of tuberculosis. *International Journal of Tuberculosis and Lung Disease* 4, 550–554.