Validity of Self-Reported Hormone Replacement Therapy Use in Manitoba

Erich Kliewer

Teresa Mayer

Alain Demers

Marion Harrison

Final Report
Contract No. OSC3090

July, 2001

CancerCare MANITOBA

Table of Contents

List of Tables and Figures

1. Introduction

Hormone replacement therapy (HRT) generally refers to the use of estrogen or it analogues with or without progestin (1). It is commonly prescribed to women in order to alleviate symptoms associated with menopause and for long-term prevention of osteoporosis (2), and potentially cardiovascular disease (3) and Alzheimer's disease (4,5). In Canada, 21% of women between the ages of 50 and 69 reported that they currently used HRT (6).

HRT use has been associated with an increased risk of endometrial and breast cancer (7). Use of estrogen in combination with a progestin reduces the risk of endometrial cancer considerably, although it does not eliminate it (7,8). On the other hand, there are indications that the combined therapy has a greater impact on breast cancer risk than estrogen alone (9-13). HRT use has other impacts on medical outcomes. It has been reported that HRT use reduces the sensitivity of mammographic screening because it leads to an increase in breast density (18-20). This reduction in sensitivity may impact population-based screening programs by increasing the number of missed cancers (false negative exams) and ultimately reduce the potential mortality benefit from screening. Harvey et al., have found that stopping HRT use for two weeks can result in a regression of hormone-induced changes in breast density (15). Thus, knowledge of a woman's current use of HRT, rather than past use, may be more informative to the radiologist in interpreting the film and in making a more accurate diagnosis.

Organized breast cancer screening programs in Canada rely on self-report by women to determine if they are currently using HRT. The assumption is made by the screening programs that self-report of HRT use is relatively accurate, but to our knowledge no validation studies of HRT use have been undertaken in Canada. Although several studies have examined the validity or reliability of self-reported HRT use in other countries, most have focused on prior use (21-28) rather than current use (29). These studies have generally found self-reported HRT use is fairly accurate, although it tends to vary by the demographic and socioeconomic characteristics of the woman (24,27-29), as well as by use of other drugs (29), time between the interview and the date that the drug was last dispensed (26), duration of use (27), and route of administration (27).

The purpose of this study was to examine the validity of self-reported HRT use among various subgroups of Manitoba women 40 years of age and over. Self-reported information from a sample of the general population of Manitoba women from the National Population Health Survey (NPHS) and from women attending the Manitoba Breast Screening Program (MBSP) were compared with prescription drug use information recorded in Manitoba Health's Drug Prescription Information Network (DPIN).

2. Methods

2.1 Data sources

In this study, self-reported HRT information for women residing in Manitoba was compared to prescription drug information from the Manitoba Health DPIN database. Two sources of self-reported HRT use were used: the 1996/97 NPHS and the MBSP questionnaire for the years 1995-99.

2.1.1 National Population Health Survey (NPHS)

In 1996/97, the second round of the NPHS was administered by Statistics Canada to a large sample of residents of Canada who were chosen through a multi-stage cluster design. The NPHS was composed of two sections: (1) a general or household survey asked of all household members; and (2) a health section, in which only one family member aged 12 years or older was selected. Data used in this report comes from those individuals who answered the health portion of the survey. First Nations communities and persons residing on Canadian Force Bases were not included in the sample for the NPHS.

The NPHS asked respondents if they were willing to have their data linked to health records for research purposes (30).[1] A special share file for the 1996/97 NPHS was created for individuals who agreed to have their questionnaires linked to health records for research and this file was used for our analyses. The share file contains the Personal Health Identification Number (PHIN). Many respondents supplied a PHIN, but for those respondents who provided a Manitoba Health Services Insurance Plan (MHSIP) Registration Number or incorrect PHIN, the PHIN was determined by linking the NPHS file to the Manitoba Health Population Registry (MHPR). Probabilistic linkage, based on name, birth date, sex, and in some cases MHSIP Registration Number, was used.

For the 1996/97 NPHS, the Manitoba Government paid Statistics Canada to increase the Manitoba sample size to enable more detailed research. The survey was completed by 11,431 residents of Manitoba, of whom 3,573 were women 40 years of age and over. All women aged 30 and over were asked if they had used HRT. Of the 3,573 women aged 40 and over, 3,465 (97.0%) stated that their survey data could be linked for research purposes. Of these women, 3,233 had a successful match to a PHIN. A further 8 women did not answer the question, “In the past month, did you take hormones for menopause or aging symptoms?” with a valid response

[1] The wording in the NPHS on linking survey data with administrative data is as follows:

LINK-INT
We are seeking your permission to link information collected during this interview with provincial health information. This would include information on past and continuing use of services such as visits to hospitals, clinics, doctor's offices or other services provided by the province. Do we have your permission?

AM66-LNK
1 Yes
2 No

(yes or no) and were excluded from the study. This final sample size (N=3,225) was 90.3% of all NPHS respondents who were women aged 40 years and over and residents of Manitoba. Table 1 shows the frequency distributions of the characteristics of the women who were included and excluded from the analyses.

Table 1. Characteristics of women in the NPHS who were included and excluded in the analyses

Characteristics		Included		Excluded	
		N	%	N	%
Age	40-49	868	26.9	82	23.6
	50-59	700	21.7	81	23.3
	60-69	657	20.4	55	15.8
	70+	1,000	31.0	130	37.4
Mean Age		61.0	13.9[1]	62.9	14.9[1]
Region of residence	Rural	1,316	40.8	130	37.4
	Urban	1,909	59.2	218	62.6
Marital status	Single	156	4.8	25	7.2
	Widow	949	29.4	119	34.2
	Divorced/separated	274	8.5	33	9.5
	Married/common-law/partner	1,838	57.0	167	48.0
	Missing	8	0.3	4	1.1
Country of birth	Canada	2,867	88.9	307	88.2
	Not Canada	357	11.1	37	10.6
	Unknown	1	0.0	4	1.1
Ethnicity	Aboriginal	125	3.9	16	4.6
	Non-Aboriginal	3,100	96.1	332	95.4
Languages spoken	English & Other	1,107	34.3	156	44.8
	English only	2,097	65.0	191	54.9
	No English	20	0.6	1	0.3
	Not stated	1	0.0	--	--
Education	< Secondary graduation	1,386	43.0	175	50.3
	>= High school graduation	1,811	56.2	162	46.6
	Missing	28	0.9	11	3.2
Family income	No income or < $30,000	1,463	45.4	157	45.1
	$30,000 or more	1,230	38.1	83	23.9
	Not stated	532	16.5	108	31.0
Occupation	Administration/science/arts/religion	536	16.6	58	16.7
	Clerical/sales/services	627	19.4	53	15.2
	Farming/fishing/forestry	56	1.7	4	1.1
	Processing/construction/crafts	79	2.5	10	2.9
	Not applicable	1,904	59.0	216	62.1
	Not stated	23	0.7	7	2.0

1. Standard deviation.

Age was calculated at the time of the last NPHS interview date, using the Manitoba Health Population Registry (MHPR) birth date. Urban/Rural status was determined from the NPHS variable "GE36DURB". Urban includes Winnipeg and Brandon. Aboriginal ethnic origin indicates whether the respondent stated that their ethnic or cultural group was either (1) North American Indian, (2) Metis or (3) Inuit/Eskimo and Native status was based on responses that

race or colour was (1) Native, (2) Aboriginal peoples of North America or (3) North American Indian, Metis or Inuit/Eskimo.

Compared to the women included in the study, among the excluded NPHS respondents there was:

- a lower percentage of women aged 60-69 and a higher percentage of women aged 70+
- a lower percentage of women who were married or living with a partner
- a higher percentage of women who spoke English and another language
- a higher percentage of women with less than secondary school graduation
- a lower percentage of women with an income of $30,000 or more.

2.1.2 Manitoba Breast Screening Program (MBSP)

The MBSP administers a questionnaire to all participants to collect information about demographics, breast cancer risk factors, and relevant medical history. To obtain information on hormone use first time screeners are asked “Have you ever taken estrogen (hormones for menopause)?" Valid responses are “Not sure,” “No” or “Yes.” If respondents answered yes, they were then asked, "Are you currently taking them now?” Women with invalid responses (missing, spoiled or no response) to the question “Have you ever taken estrogen (hormones for menopause)?” were excluded. Women who had previously attended (prevalent screens) were asked, “Since your last screening visit have you taken hormones?” Valid responses are “No” or “Yes.”

The number of first-time screeners aged 50-69 in the years 1995 to 1999 was 59,616. Among these women, 294 had invalid, missing or spoiled responses to the question, “Have you ever taken estrogen (hormones for menopause)?” An additional 564 women did not respond to the question, “If yes, are you currently taking [HRT] now?” Of the remaining women, seven did not successfully link to the MHPR file (PHIN could not be validated.) These 865 women were excluded, resulting in a total of 58,751 (98.5%) women in the study. An additional 1,639 women were excluded because they were not in the 50-69 year age group. However, women under the age of 50 (N=733) and over the age of 70 (N=906) were included in an age group stratification analysis.

Compared to the women included in the study, among the excluded first-time screeners there was:

- a greater percentage of women aged 60-69
- a lower percentage of women of British ethnicity, but a greater percentage of women with ‘Missing’ ethnicity
- a lower percentage with high school graduation, but a greater percentage of women with ‘Missing’ education
- a greater percentage of women who had their mammogram in 1997 and a lower percentage who had it in 1998 and 1999.

Table 2. Characteristics of incident screeners in the MBSP who were included and excluded in the analyses

Characteristics		Included		Excluded	
		N	%	N	%
Age[1]	40-49	733	1.2	--	
	50-59	35,441	58.7	476	55.0
	60-69	23,310	38.6	389	45.0
	70+	906	1.5	--	
Mean Age		58.0	6.3[2]	58.8	5.9[2]
Region of residence	Rural	24,094	41.0	347	40.1
	Urban	34,583	58.9	517	59.8
	Missing	74	0.1	1	0.1
Ethnic background	Canadian	3,802	6.5	53	6.1
	Aboriginal	1,841	3.1	28	3.2
	Asian	1,812	3.1	33	3.8
	British	19,646	33.4	222	25.7
	French	4,693	8.0	60	6.9
	Western European	9,986	17.0	165	19.1
	North European	2,293	3.9	28	3.2
	Eastern European	9,383	16.0	122	14.1
	South European	1,191	2.0	32	3.7
	Other European	1,025	1.7	25	2.9
	Other	681	1.2	12	1.4
	Missing	2,398	4.1	85	9.8
Education					
	Some High School or less	25,038	42.6	400	46.2
	High School graduation	21,201	36.1	265	30.6
	Degree or Diploma	11,948	20.3	149	17.2
	Missing	564	1.0	51	5.9
Year					
	1995	2,580	4.4	47	5.4
	1996	12,698	21.6	214	24.7
	1997	16,644	28.3	317	36.7
	1998	13,906	23.7	167	19.3
	1999	12,923	22.0	120	13.9

1. Women under 50 years of age and over 69 years of age were excluded from all other frequency tabulations.
2. Standard deviation.

During the period 1995 through 1999 there were 25,651 prevalent screeners in the MBSP aged 50-69 years (Table 3). Of these women, 255 had an invalid, missing or spoiled response to the question pertaining to taking hormones since the last mammogram screen. These women were excluded, resulting in 25,396 women (99.0%) in the analyses. It should be noted that since prevalent screeners were not specifically asked about current use, but rather, use since their last mammogram, they may not be current users.

Compared to the women included in the study, among the excluded prevalent screeners there was:

- a greater percentage of rural women

- a greater percentage of women who were screened in 1996 and 1997, and a lower percentage of women who were screened in 1999.

Table 3. Characteristics of prevalent screeners in the MBSP who were included and excluded in the analyses

Characteristics		Included		Excluded	
		N	%	N	%
Age	50-59	12,420	48.9	133	52.2
	60-69	12,976	51.1	122	47.8
Mean Age		59.9	5.3[1]	59.5	5.4[1]
Region of residence	Rural	10,138	39.9	135	52.9
	Urban	15,230	60.0	120	47.1
	Missing	28	0.1	--	--
Ethnic background	Aboriginal	467	1.8	7	2.8
	Other	305	1.2	2	0.8
	Asian	515	2.0	7	2.8
	British	9,520	37.5	90	35.3
	French	1,964	7.7	23	9.0
	Western European	4,181	16.5	47	18.4
	North European	1,117	4.4	10	3.9
	Eastern European	4,089	16.1	34	13.3
	South European	459	1.8	2	0.8
	Other European	449	1.8	1	0.4
	Canadian	1,389	5.5	19	7.5
	Missing	941	3.7	13	5.1
Education	Some High School or less	10,445	41.1	115	45.1
	High School graduation	9,908	39.0	91	35.7
	Degree or Diploma	4,854	19.1	47	18.4
	Missing	189	0.7	2	0.8
Year	1995			1	0.4
	1996	128	0.5	21	8.2
	1997	1,967	7.8	40	15.7
	1998	8,994	35.4	101	39.6
	1999	14,307	56.3	92	36.1
Screen/Visit number	2	23,173	91.3	240	94.1
	3	2,135	8.4	15	5.9
	4	83	0.3	--	--
	5	5	0.0	--	--

1. Standard deviation.

The key variables available for the analysis of the MBSP data were:

Age. Age was calculated using birth date at the time of the MBSP screen. Since the MBSP is directed at women aged 50-69, women were grouped into age categories of 50-59 and 60-69. A total of 1,639 women under 50 years of age and 70 years of age or older were included in the age distribution shown in Table 2 and in one analysis table (Table 10).

Urban/Rural Residence. Urban/rural residence was based on a woman's residential postal code and is verified at the time of the screen. Urban residence was defined as having a Winnipeg or Brandon postal code.

Ethnic Background. Women were asked, "To what ethnic or cultural group did your ancestors belong?" For the analysis, ethnic background was grouped into (1) Aboriginal, (2) Asian, (3) British, (4) French, (5) Western European, (6) Northern European, (7) Eastern European, (8) Southern European, (9) Canadian and (10) Other (African/Caribbean/Arabian/Pacific Islands /South and Central America/Other).

Education. Women were asked during their first screen what their highest level of education was ("How far did you go in school?") Education level was grouped as: (1) Some high school / Grade 9 or less, (2) High school graduation (which includes some college/university), or (3) Degree or diploma.

2.1.3 Drug Prescription Information Network (DPIN)

The DPIN is an administrative database maintained by Manitoba Health that contains dispensing information on prescriptions from all community-based pharmacies and is networked throughout the province. All prescriptions submitted by retail pharmacies for drug insurance reimbursement or drug utilization review are contained in this file. Reimbursement is provided for prescriptions dispensed to residents of Manitoba in the provincial Pharmacare program. Most residents receive partial funding for pharmaceuticals through the Pharmacare program, so most prescriptions are recorded on the DPIN file. Prescriptions dispensed to patients while in hospital and to Status Indians through nursing stations are excluded, although most outpatient hospital prescriptions are included.

The DPIN was established in 1995 and includes patient information such as the PHIN, Manitoba Health registration number, birth date, sex and patient postal code; drug information such as the drug dispensed (which is a 9-digit drug identification number (DIN)), the date dispensed, number of days supply, metric quantity of drug dispensed, and other information such as physician and pharmacy identifiers, and costs. The DPIN database also contains a file (the master formulary) with all possible DINs. The master formulary contains specific drug information for each DIN, such as product name, therapeutic class (e.g. estrogens), dosage form (e.g. tablet, intravenous) and strength (31-33).

For this study, hormone replacement therapy was identified through the master formulary, first by identifying estrogens and progestins. A pharmacist then went through the list of drugs and identified which medications were primarily prescribed for hormone replacement therapy. The DINs obtained from this list were then abstracted from the DPIN prescription database. The list of drugs included in the study are outlined in Table 4.

Table 4. List of possible Hormone Replacement Therapies[1]

Brand name	Dosage	Formulation	Strength	DIN
Apo-Conest	Tablet	Conjugated estrogen	0.3mg	00798231
			0.625mg	00798223
			0.9mg	00798215
			1.25mg	00798207
			2.5mg	00798193
C.E.S.	Tablet	Conjugated estrogen	0.3mg	02230891
			0.9mg	02230892
			0.3mg	00486574
			0.9mg	00831395
			0.625mg	00265470
			1.25mg	00265489
Climara	Patch	17β-estradiol	3.9mg	02231509
			7.8mg	02231510
Congest	Tablet	Conjugated estrogen	0.9mg	00820224
			0.3mg	00830240
			0.625mg	00830232
			0.9mg	00830224
			1.25mg	00830216
			2.5mg	00830208
Conjugated estrogens	Tablet		0.625	00587281
			1.25	00587303
Estinyl	Tablet	Ethinyl estradiol	0.05mg	00028223
			0.5mg	00028231
			0.02mg	00028215
Estrace	Tablet	micronized 17β-estradiol	0.5mg	02225190
			1mg	00464791
			1mg	02148587
			2mg	00464805
			2mg	02148595
	Vaginal cream	micronized 17β-estradiol		
Estracomb	Patch	17β-estradiol / norethindrone acetate		02108186
Estraderm	Patch	17β-estradiol	100	00756792
			25	00756849
			50	00756857
Estring	Vaginal ring	17β-estradiol		02168898
Estrogel transdermal	Gel	17β-estradiol	0.06%	02238704
Gen-Medroxy	Tablet	Medroxyprogresterone acetate	10mg	02229840
Kenral-MPA (Alti-MPA)	Tablet	Medroxyprogresterone acetate	2.5mg	02148552
			5mg	02148560
			10mg	02148579
Neo-Estrone	Tablet	Esterified estrogens	0.625	00473618
			1.25mg	00287725
			1mg	00287733
	Vaginal cream	Estrone	1mg	00727369
Novo-Medrone	Tablet	Medroxyprogresterone acetate	2.5mg	02221284
			5mg	02221292
			10mg	02221306
Oesclim	Patch	17β-estradiol	10mcg	02237807
			50mcg	02237808
Oestrilin	Vaginal cones	Estrone		00006211
	Vaginal cream	Estrone	1mg	00006149

Table 4. List of possible Hormone Replacement Therapies[1] (continued)

Brand name	Dosage	Formulation	Strength	DIN
Ogen 0.625	Tablet	Estropipate	0.75mg	02016958
Ogen 0.625		Estropipate	0.75mg	02089793
Ogen 1.25		Estropipate	1.5mg	02089769
Ogen 2.5		Estropipate	3.0mg	02089777
Ortho Dinestrol	Vaginal cream	dinestrol	0.1mg/g	00441295
			0.1mg/g	00990531
Premarin	Tablet	Conjugated estrogen	0.625mg	02043408
			2.5mg	02043432
			0.5ml	00002569
			0.3mg	02043394
			0.625mg	00002577
			0.9mg	00403466
			0.9mg	02043416
			1.25mg	00002585
			1.25mg	02043424
			2.5mg	00002593
Premarin methyltestost	Tablet		0.625mg	00053538
Premarin	Vaginal cream	Conjugated estrogen	0.625mg/gm	00002089
			0.625mg	02043440
Proclim	Tablet	Medroxyprogesterone acetate	2.5mg	02239825
			5mg	02239826
			10mg	02239827
Prometrium	Cap	Progesterone	100mg	02166704
Provera	Tablet	Medroxyprogesterone acetate	5mg	00030937
			2.5mg	00708917
			10mg	00729973
	Tablet-U		10mg	02010933
	Tablet-W		5mg	02010739
Vivelle	Patch	17β-estradiol	100ug	02204444
			37.5ug	02204401
			50ug	02204428
			75ug	02204436

1. Source: The Compendium of Pharmaceutical Specialities (34) and the DPIN master formulary.

2.1.4 Manitoba Health Population Registry (MHPR)

The MHPR contains the registration information for all individuals who are eligible for Manitoba medical insurance. Information in this file was used to verify the PHIN for women in the NPHS and MBSP files. As well, the MHPR was used to determine the number of Manitoba women aged 40 years and over. Using this population, the percentages of women included in the NPHS and MBSP were calculated (Table 5). Of Manitoba women aged 50-69, 1.33% were included in the NPHS and 12.94% were included in the MBSP. Note, that the latter represents the average percentage of incident screeners over the five-year period 1995-99.

Table 5. Percentage of Manitoba women included in NPHS and MBSP survey, by age

Age	1996/97 NPHS			1995-99 MBSP first time screeners		
	N	Population[1]	%	N	Population[2]	%[3]
40-49	868	83,289	1.04	733	376,736	0.19
50-59	700	57,628	1.21	35,441	261,492	13.55
60-69	657	44,477	1.48	23,310	200,159	11.65
70+	1,000	67,898	1.47	906	306,271	0.30
Total 40+	3,225	253,292	1.27	60,390	1,144,658	5.28
Total 50-69	1,357	102,105	1.33	59,751	461,651	12.94

1. Source: MHPR - population as at December 31, 1997.
2. Source: MHPR - population is the sum of the 1995-99 December 31 populations. The population in 1995 was adjusted to take into account that the MBSP did not start until July of that year.
3. Average percentage of incident screeners over the five-year period 1995-99.

2.2 Data linkage

All files described above have a unique common identifier, the PHIN. However, the DPIN file is not always accurate in reporting the correct PHIN. Because of this, the Manitoba Health registration number was obtained from the MHPR for all women in the NPHS and MBSP. DPIN records were then selected if there was an HRT DIN, sex was female and the MH registration number matched that from the NPHS and MBSP files. Since the registration number is a family identifier, another step was taken to ensure the retrieved record was specific to the women in the surveys. Thus, after the DPIN data was retrieved, the PHIN or birth date from the DPIN data was matched to the PHIN or birthdate from the MHPR or survey file. If there was an exact match on either PHIN or birthdate, then the record was kept. Those women removed from the file had an unmatched PHIN and an unmatched birth date. The MHPR PHIN was added to the DPIN data for later linkage.

Women who satisfied the inclusion criteria for the DPIN extract were then linked to the NPHS and MBSP files and examined for current HRT use. The dispense date and days supply from the DPIN file were used to validate whether women were current users or previous users. A 'prescription end date' in the DPIN file was calculated by adding the number of days supply plus 10 to the dispense date. The 10 days were added in order to allow for some flexibility in the reporting of current use. In the NPHS, current HRT use was defined as currently using HRT or using HRT within the 31 days prior to the interview. In order to confirm current use we determined if the NPHS interview date or any time in the 31 days prior to the interview occurred between the DPIN HRT dispense date and the 'prescription end date'. For the MBSP incident screeners, current HRT use was defined as using HRT on the date of the mammography screen. Thus, if the screen date fell between the dispense date and the 'prescription end date', the woman was confirmed as currently using HRT. Among the MBSP prevalent screeners, HRT utilization was defined as using HRT at any time since the previous MBSP screen. Thus, HRT use was confirmed if the previous screen date and the current screen date fell between an HRT dispense date and the "prescription end date." Women who did not have a current HRT prescription according to the DPIN file were flagged as not being a current HRT user. Evidence of HRT utilization on the DPIN file was considered to be the "gold standard."

2.3. Analysis

2.3.1 Validity measures

In order to determine the validity of self-reported hormone replacement therapy, the concordance, sensitivity, positive predictive value (PPV), negative predictive value (NPV) and kappa statistic were calculated. Figure 1 outlines how each of the measures were calculated.

Figure 1. Self-reported current HRT use compared to current DPIN HRT prescriptions.

		DPIN: Yes	DPIN: No
Self-report	Yes	a	b
	No	c	d

Concordance = ((a + d) / (a + b + c + d)) * 100
This is the percentage of women for whom the self-report and the DPIN data agree.

Sensitivity = (a / (a + c)) * 100
This is the percentage of women currently using HRT who accurately report their use.

Specificity = (d / (b + d)) * 100
This is the percentage of women currently not using HRT who accurately report they are not using it.

Positive Predictive Value = (a / (a + b)) * 100
This is the percentage of women who self-reported currently using HRT and are in fact currently using HRT.

Negative Predictive Value = (d / (c + d)) * 100
This is the percentage of women who self-reported currently not using HRT and are in fact currently not using HRT.

The kappa statistic is a measure of nonrandom agreement between two measurements of the same categorical variable. The kappa statistic tests the ability to replicate the information whether or not the information is good. Although the kappa statistic is generally not used when there is a 'gold standard', we have reported it. Agreement for the kappa statistic is considered poor if it is less than 0.00, slight if it is 0.00-0.20, fair if it is 0.21-0.40, moderate if it 0.41-0.60, substantial if it is 0.61-0.80, and almost perfect if it is 0.81-1.00 (35).

2.3.2 Validity analyses

Two sets of validity analyses were performed between the NPHS and MBSP survey data and the DPIN administrative data. In the first set of analyses, overall validity measures were calculated. In the second set of analyses the validity measures were stratified by demographic and socioeconomic characteristics. Analyses of the MBSP data were stratified by age, region of residence, ethnic background and education. Analyses for the NPHS data were stratified according to these variables, and in addition, by marital status, main language spoken, region of birth, family income and occupation. Agreement on current HRT use was determined by comparing the responses in the NPHS and MBSP to information obtained from DPIN.

To test for differences in the validity measures between the levels of the stratification variables, the Freeman-Tukey test with bootstrap resampling was used. The Freeman-Tukey test is based on the double-arcsine transformed proportions (36). Bootstrap resampling was used to adjust the p-values for the multiple testing problem (37,38). Adjustments were based on resampling the data with replacement.

Although not the focus of this report, we also undertook a reliability analysis that compared the responses of women who were in both the NPHS and the MBSP.

3. Results

3.1 HRT use in Manitoba

A substantially greater percentage of MBSP first time screeners stated they were currently using HRT than women in the NPHS (Table 6). This pattern was consistent across each age group. Among women in the screening age group (50-69) only 17.5% of women in the NPHS reported they were currently using HRT, compared to 27.7% of women in the MBSP who were undergoing their first screen. It is not clear why such a large difference exists, but variations in the questions regarding HRT use may have contributed to the differences.

Table 6. Percentage of women self-reporting current HRT use in NPHS and MBSP first time screeners

Age	1996/97 NPHS[1,2]			1995-99 MBSP first time screeners[2]		
	Self-reported HRT users	N	%	Self-reported HRT users	N	%
<50	86	868	9.9	128	733	17.4
50-59	172	700	24.6	11,945	35,441	33.7
60-69	66	657	10.1	4,347	23,310	18.6
70+	23	1,000	2.3	107	906	11.8
All ages	347	3,225	10.8	16,527	60,390	27.4
Total 50-69	238	1,357	17.5	16,292	58,751	27.7

1. Based on unweighted sample.
2. As shown in Tables 1 and 2, some women have been excluded from the analysis.

In Newfoundland, 39% of women attending the breast screening program reported current HRT use while in Alberta approximately a third of screened women reported such use (personal

communication). Among women aged 50-64 attending the United Kingdom National Health Service Breast Screeening Programme, 33% reported that they were currently using HRT (39).

3.2 Comparison of self-report in the MBSP and the NPHS

As the PHIN was recorded in both the MBSP and NPHS, it was possible to compare the responses of women who were included in both to the questions regarding HRT use. The two survey files were linked by PHIN. There were 919 MBSP first-time screeners who linked to the NPHS. When the link was restricted to first-time screeners in the years 1996 or 1997, the years that the NPHS was conducted, there were 479 matches. The agreement measures for the two sets of comparisons are given in Table 7 (all NPHS-MBSP links) and Table 8 (NPHS links with 1996-97 MBSP screeners.)

Table 7. Comparison of self-reported current HRT use in the MBSP first time screeners and in the NPHS

	NPHS								
MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Yes	168	68	236	89.4	85.3	90.6	71.2	95.8	0.71
No	29	654	683	--	--	--	--	--	--
Total	197	722	919	--	--	--	--	--	--

Table 8. Comparison of self-reported current HRT use in the MBSP first time screeners (1996-97) and in the NPHS

	NPHS								
MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Yes	106	21	127	93.1	89.8	94.2	83.5	96.6	0.82
No	12	340	352	--	--	--	--	--	--
Total	118	361	479	--	--	--	--	--	--

As is to be expected, restricting the MBSP data to the time most relevant to the NPHS survey, resulted in a greater agreement. The concordance for the restricted analysis was 93.1% and the kappa was 0.82. Thus, there was good agreement between the two surveys and the reliability of self-reported current HRT use was high. Greendale et al. (27) found a concordance of 95.2% and a kappa of 0.92 for women reporting ever use of HRT in surveys that were three months apart.

Several factors may have contributed to the discrepancies in women's answers. One is the time element; it is possible that a woman was or was not taking HRT at the time of one survey, but at the time of the second survey, they were no longer using HRT or had started on HRT. Another reason for discrepancies may be due to the wording of the questions. The NPHS asked women if they had taken hormones in the last month, whereas the MBSP asked if they were currently taking hormones for menopause. For the MBSP questionnaire, a nurse goes over many of the questions with each woman, and therefore the answers on the MBSP may be more reliable than those in the NPHS.

3.3 Validity of survey data

3.3.1 Total

With the exception of sensitivity, the validity measures for current self-reported HRT use were fairly similar in both the MBSP and NPHS surveys (Tables 9,10). The sensitivity among MBSP women was 94.4%, which was substantially higher than among NPHS women (82.1%). Thus a larger proportion of women in the NPHS than in the MBSP reported they were not current users of HRT when in fact the DPIN data indicated they were. However, the NPHS sample included a large proportion of women over the age of 69, and as seen in the next section, these women were the least accurate in reporting their current HRT use.

Although the concordance and kappa were similar between first time and prevalent screeners, the sensitivity and NPV were higher among first time screeners, while specificity and PPV were lower (Tables 9,11).

Table 9. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions

	DPIN								
MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Yes	13,363	2,929	16,292	93.7	94.4	93.4	82.0	98.1	0.84
No	790	41,669	42,459	--	--	--	--	--	--
Total	14,153	44,598	58,751	--	--	--	--	--	--

Table 10. Self-reported current HRT use in the NPHS compared to DPIN prescriptions

	DPIN								
NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Yes	276	71	347	95.9	82.1	97.5	79.5	97.9	0.79
No	60	2,818	2,878	--	--	--	--	--	--
Total	336	2,889	3,225	--	--	--	--	--	--

Table 11. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions

	DPIN								
MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Yes	9,044	420	9,464	92.4	85.7	97.2	95.6	90.6	0.84
No	1,506	14,426	15,932	--	--	--	--	--	--
Total	10,550	14,846	25,396	--	--	--	--	--	--

3.3.2 Age

Among first time screeners attending the MBSP, the sensitivity and PPV were substantially lower for women 70 years of age and over than for younger women (Table 12). There was a clear pattern of decreasing sensitivity with increasing age. In contrast, although the differences were not large, concordance and specificity were highest among the elderly women. Similar patterns were observed among the NPHS women (Table 13). With the exception of sensitivity,

the validity measures of the MBSP and NPHS women were comparable. The sensitivity was substantially lower among women in the NPHS, particularly for women aged 70 years and over, where it was approximately half that observed for the MBSP women. Decreasing sensitivity with age was also observed among prevalent screeners (Table 14).

Table 12. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions, by age

Age Group	MBSP	DPIN Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
<50	Yes	86	42	128	93.9	96.6	93.5	67.2	99.5	0.76
	No	3	602	605	--	--	--	--	--	--
50-59	Yes	9,912	2,033	11,945	92.9	95.4	91.9	83.0	98.0	0.84
	No	479	23,017	23,496	--	--	--	--	--	--
60-69	Yes	3,451	896	4,347	94.8	91.7	95.4	79.4	98.4	0.82
	No	311	18,652	18,963	--	--	--	--	--	--
70+	Yes	45	31	76	95.7	84.9	96.4	59.2	99.0	0.68
	No	8	822	830	--	--	--	--	--	--
Total		14,295	46,095	60,390	--	--	--	--	--	--

Statistical differences exist between:
(a) Ages <50 and 50-59 for PPV ($p<0.0005$) and NPV ($p<0.05$).
(b) Ages <50 and 60-69 for PPV ($p<0.05$).
(c) Ages 50-59 and 60-69 for concordance ($p<0.0001$), sensitivity ($p<0.0001$), specificity ($p<0.0001$), PPV ($p<0.0001$) and NPV ($p<0.05$).
(d) Ages 50-59 and 70+ for concordance ($p<0.01$), specificity ($p<0.0001$) and PPV ($p<0.0005$).
(e) Ages 60-69 and 70+ for PPV ($p<0.005$).

Table 13. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by age

Age Group	NPHS	DPIN Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
40-49	Yes	60	26	86	96.2	89.6	96.8	69.8	99.1	0.76
	No	7	775	782	--	--	--	--	--	--
50-59	Yes	144	28	172	92.9	86.8	94.8	83.7	95.8	0.81
	No	22	506	528	--	--	--	--	--	--
60-69	Yes	55	11	66	97.0	85.9	98.2	83.3	98.5	0.84
	No	9	582	591	--	--	--	--	--	--
70+	Yes	17	6	23	97.2	43.6	99.4	73.9	97.8	0.53
	No	22	955	977	--	--	--	--	--	--
Total		336	2,889	3,225	--	--	--	--	--	--

Statistical differences exist between:
(a) Ages <50 and 50-59 for NPV ($p<0.005$).
(b) Ages <50 and 70+ for sensitivity ($p<0.0005$) and specificity ($p<0.005$).
(c) Ages 50-59 and 60-69 for concordance ($p<0.05$) and specificity ($p<0.05$).
(d) Ages 50-59 and 70+ for concordance ($p<0.005$), sensitivity ($p<0.0001$) and specificity ($p<0.0001$).

As was observed for the overall results (Section 3.2.1), for women aged 50-69 years the sensitivity and NPV were higher among first time screeners than prevalent screeners, while the reverse was true for specificity and PPV (Tables 12,14).

Table 14. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions, by age

Age		DPIN								
Group	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
50-59	Yes	5,521	247	5,768	92.5	88.9	96.0	95.7	89.6	0.85
	No	690	5,962	6,652	--	--	--	--	--	--
60-69	Yes	3,523	173	3,696	92.4	81.2	98.0	95.3	91.2	0.82
	No	816	8,464	9,280	--	--	--	--	--	--
Total		10,550	14,846	25,396	--	--	--	--	--	--

Statistical differences exist for sensitivity ($p<0.0001$), specificity ($p<0.0001$), and NPV ($p<0.005$).

Our findings differ from those of previous studies, although they did not examine current HRT use. Goodman et al., who compared self-reported ever use of menopausal estrogens with physician records, did not find any age differences in agreement (24). West et al. compared questionnaire results with information in a pharmacy database and did not find age differences in the recall accuracy of the name of the estrogen that was used (26). Although Van den Brandt examined accuracy of recall of long-term current use of numerous drugs (not specifically HRT), they did find that accuracy decreased with age (40).

3.3.3 Region of residence

There were only minor differences in the validity measures between urban and rural residents, although some of the differences reached statistical significance due to the large number of women included in the analyses (Tables 15-17).

Table 15. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions, by region of residence

		DPIN								
Region	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Rural	Yes	4,874	1,135	6,009	94.0	94.1	94.0	81.1	98.3	0.83
	No	306	17,779	18,085	--	--	--	--	--	--
Urban	Yes	8,468	1,791	10,259	93.4	94.6	93.0	82.5	98.0	0.84
	No	483	23,841	24,324	--	--	--	--	--	--
Total		14,131	44,546	58,677	--	--	--	--	--	--

Statistical differences exist for concordance ($p<0.05$) and specificity ($p<0.0001$).

Table 16. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by region of residence

		DPIN								
Region	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Rural	Yes	109	27	136	95.9	80.2	97.7	80.2	97.7	0.78
	No	27	1,153	1,180	--	--	--	--	--	--
Urban	Yes	167	44	211	96.0	83.5	97.4	79.2	98.1	0.79
	No	33	1,665	1,698	--	--	--	--	--	--
Total		336	2,889	3,225	--	--	--	--	--	--

No statistical differences exist.

Table 17. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions, by region of residence

		DPIN								
Region	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Rural	Yes	3,280	161	3,441	92.4	84.3	97.4	95.3	90.9	0.84
	No	611	6,086	6,697	--	--	--	--	--	--
Urban	Yes	5,750	258	6,008	92.4	86.6	97.0	95.7	90.3	0.84
	No	893	8,329	9,222	--	--	--	--	--	--
Total		10,534	14,834	25,368	--	--	--	--	--	--

Statistical differences exist for sensitivity ($p<0.005$).

3.3.4 Marital status

The validity of self-reported HRT use by marital groups in the NPHS showed inconsistent results (Table 18). Divorced or separated women had the lowest concordance (92.3%), specificity (92.3%), positive predictive value (56.8%) and kappa statistic (0.66), but the highest sensitivity (92.6%). Widowed women had the lowest sensitivity (64.0%), but the highest concordance (96.9%) and specificity (98.8%). Married women had the highest PPV (85.4%) and kappa statistic (0.83). Married/partnered women were the most consistent, having the highest or second highest scores on all validity measures except for NPV.

As in this study, Merlo et al. found that divorced women had the lowest accuracy in reporting current HRT use, while widowed women had the highest (29). Goodman et al., examining ever use of menopausal estrogens found no differences in concordance by marital status (24). In terms of ability to recall the name of the estrogen that was used, West et al. did find that in univariate analyses married women had better recall, but that this difference disappeared when education and smoking status were considered (28).

Table 18. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by marital status

Marital Status	NPHS	DPIN Yes	DPIN No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Married / Partner	Yes	210	36	246	96.0	84.7	97.7	85.4	97.6	0.83
	No	38	1,554	1,592	--	--	--	--	--	--
Single	Yes	9	5	14	95.5	81.8	96.6	64.3	98.6	0.70
	No	2	140	142	--	--	--	--	--	--
Widowed	Yes	32	11	43	96.9	64.0	98.8	74.4	98.0	0.67
	No	18	888	906	--	--	--	--	--	--
Divorced / Separated	Yes	25	19	44	92.3	92.6	92.3	56.8	99.1	0.66
	No	2	228	230	--	--	--	--	--	--
Total		336	2,881	3,217	--	--	--	--	--	--

Statistical differences exist between:
(a) Married/Common-law and Widowed for sensitivity ($p<0.05$)
(b) Married/Common-law and Divorced/Separated for specificity ($p<0.005$) and PPV ($p<0.005$)
(c) Widowed and Divorced/Separated for concordance ($p<0.05$) and specificity ($p<0.0001$).

3.3.5 Birthplace and ethnicity

In the NPHS all of the validity measures were higher for native-born Canadians than the foreign-born (Table 19), although none of the differences were statistically significant.

Table 19. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by birthplace

Country	NPHS	DPIN Yes	DPIN No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Canada	Yes	245	59	304	96.1	82.2	97.7	80.6	97.9	0.79
	No	53	2,510	2,563	--	--	--	--	--	--
Other	Yes	31	12	43	94.7	81.6	96.2	72.1	97.8	0.74
	No	7	307	314	--	--	--	--	--	--
Total		336	2,888	3,224	--	--	--	--	--	--

No statistical differences exist.

Among both incident and prevalent MBSP screeners, Aboriginal women had lower values for all validity measures than non-Aboriginal women (Tables 20,21). Among incident screeners the differences were statistically significant for concordance, sensitivity and PPV, but among prevalent screeners significant differences only existed for sensitivity.

Table 20. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions, by Aboriginal ethnic background

		DPIN								
Ethnicity	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Aboriginal	Yes	256	111	367	91.9	87.1	92.8	69.8	97.4	0.73
	No	38	1,436	1,474	--	--	--	--	--	--
Non-Aboriginal	Yes	12,586	2,694	15,280	93.8	94.7	93.5	82.4	98.2	0.84
	No	704	38,528	39,232	--	--	--	--	--	--
Total		13,584	42,769	56,353	--	--	--	--	--	--

Statistical differences exist for concordance ($p<0.01$), sensitivity ($p<0.0001$) and PPV ($p<0.0001$).

Table 21. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions, by Aboriginal ethnic background

		DPIN								
Ethnicity	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Aboriginal	Yes	134	10	144	89.9	78.4	96.6	93.1	88.5	0.78
	No	37	286	323	--	--	--	--	--	--
Non-Aboriginal	Yes	8,594	385	8,979	92.5	85.9	97.2	95.7	90.6	0.85
	No	1,413	13,596	15,009	--	--	--	--	--	--
Total		10,178	14,277	24,455	--	--	--	--	--	--

Statistical differences exist for sensitivity ($p<0.05$).

In the NPHS, with the exception of specificity, the validity measures were higher in the Aboriginal than in the non-Aboriginal population, although the differences were not significant (Table 22). Aboriginal status could also be determined in the NPHS through a question on race. The results for race have not been presented but are similar to those based on the ethnicity question. It should be noted that the number of Aboriginal women in the NPHS was small and included only those living outside of First Nations communities. This may have contributed to the differences in the results between the MBSP and NPHS.

Table 22. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by Aboriginal ethnic background

		DPIN								
Ethnicity	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Aboriginal	Yes	17	4	21	96.0	94.4	96.3	81.0	99.0	0.85
	No	1	103	104	--	--	--	--	--	--
Non-Aboriginal	Yes	259	67	326	95.9	81.4	97.6	79.4	97.9	0.78
	No	59	2,715	2,774	--	--	--	--	--	--
Total		336	2,889	3,225	--	--	--	--	--	--

No statistical differences exist.

In the MBSP there was little consistency in the validity measure across ethnic groups for both incident and prevalent screeners (Tables 23,24). Among incident MBSP screeners (Table 23), Aboriginals had the lowest concordance (91.9%), PPV (69.8%) and kappa (0.73), while Southern Europeans had the lowest sensitivity (85.9%) and NPV (96.8%). Other Europeans had the lowest

specificity (91.5%). Eastern Europeans had the highest concordance (94.4%), the French had the highest sensitivity (95.9%), and along with the British, had the highest kappa (0.85). The British also had the highest PPV (83.5%). The highest specificity was observed among the Asians (94.4%) while the "Other' had the highest NPV (98.7%). Statistically significant differences occurred primarily for sensitivity and the PPV.

Table 23. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions, by ethnic background

		DPIN								
Ethnicity	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Canadian	Yes	896	207	1,103	93.0	93.7	92.7	81.2	97.8	0.82
	No	60	2,639	2,699	--	--	--	--	--	--
Aboriginal	Yes	256	111	367	91.9	87.1	92.8	69.8	97.4	0.73
	No	38	1,436	1,474	--	--	--	--	--	--
British	Yes	4,981	987	5,968	93.7	95.3	93.2	83.5	98.2	0.85
	No	247	13,431	13,678	--	--	--	--	--	--
French	Yes	1,143	238	1,381	93.9	95.9	93.2	82.8	98.5	0.85
	No	49	3,263	3,312	--	--	--	--	--	--
N. Europe	Yes	501	125	626	93.3	94.7	92.9	80.0	98.3	0.82
	No	28	1,639	1,667	--	--	--	--	--	--
W. Europe	Yes	2,266	464	2,730	94.0	94.4	93.9	83.0	98.1	0.84
	No	135	7,121	7,256	--	--	--	--	--	--
E. Europe	Yes	1,933	426	2,359	94.4	94.9	94.2	81.9	98.5	0.84
	No	103	6,921	7,024	--	--	--	--	--	--
S. Europe	Yes	183	60	243	92.4	85.9	93.9	75.3	96.8	0.76
	No	30	918	948	--	--	--	--	--	--
Other	Yes	301	60	361	92.2	93.8	91.5	83.4	97.0	0.82
European	No	20	644	664	--	--	--	--	--	--
Asian	Yes	265	85	350	93.9	91.4	94.4	75.7	98.3	0.79
	No	25	1,437	1,462	--	--	--	--	--	--
Other	Yes	117	42	159	92.8	94.4	92.5	73.6	98.7	0.78
	No	7	515	522	--	--	--	--	--	--
Total		13,584	42,769	56,353	--	--	--	--	--	--

Statistical differences exist between:
(a) Asian and British, for PPV ($p<0.05$)
(b) British and Aboriginal, for sensitivity ($p<0.0005$) and PPV ($p<0.0001$)
(c) French and South European, for sensitivity ($p<0.001$)
(d) French and Aboriginal, for sensitivity ($p<0.0005$) and PPV ($p<0.0001$)
(e) Western European and South European, for sensitivity ($p<0.01$)
(f) Western European and Aboriginal for sensitivity ($p<0.01$) and PPV ($p<0.0001$)
(g) North European and South European, for sensitivity ($p<0.05$)
(h) North European and Aboriginal for sensitivity ($p<0.05$) and PPV ($p<0.05$)
(i) Eastern European and South European, for sensitivity ($p<0.005$)
(j) Eastern European and Aboriginal, for concordance ($p<0.05$), sensitivity ($p<0.005$) and PPV ($p<0.0005$)
(k) Other European and Aboriginal, for PPV ($p<0.005$)
(l) Canadian and Aboriginal, for PPV ($p<0.005$).

Among the prevalent MBSP screeners, there was some consistency in the validity measures across the various ethnic groups (Table 24). Aboriginals and Asians had the lowest concordance (89.9%), sensitivity (78.4%, 81.0%), PPV (93.1%, 88.1%), and kappa (0.78, 0.77). Asians also had the lowest specificity (94.4%) and the Aboriginals also had the lowest NPV (88.5%). The French and Other Europeans had the highest concordance (93.3%), sensitivity (87.6%, 91.3%) and kappa (0.86, 0.87). The French also had the highest NPV (91.6%). The British and Northern Europeans had the highest PPV (96.3%), while the Southern Europeans had the highest specificity (98.0%). However, there were only few statistically significant differences. Other Europeans had statistically higher sensitivity than Aboriginals, while the French and British had a statistically higher PPV than the Asians.

Table 24. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions, by ethnic background

		DPIN								
Ethnicity	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Canadian	Yes	467	29	496	91.7	84.3	96.5	94.2	90.3	0.82
	No	87	806	893	--	--	--	--	--	--
Aboriginal	Yes	134	10	144	89.9	78.4	96.6	93.1	88.5	0.78
	No	37	286	323	--	--	--	--	--	--
British	Yes	3,604	137	3,741	92.7	86.6	97.4	96.3	90.4	0.85
	No	557	5,222	5,779	--	--	--	--	--	--
French	Yes	719	29	748	93.3	87.6	97.5	96.1	91.6	0.86
	No	102	1,114	1,216	--	--	--	--	--	--
N. Europe	Yes	386	15	401	92.8	85.4	97.7	96.3	90.8	0.85
	No	66	650	716	--	--	--	--	--	--
W. Europe	Yes	1,493	67	1,560	92.3	85.4	97.3	95.7	90.3	0.84
	No	255	2,366	2,621	--	--	--	--	--	--
E. Europe	Yes	1,337	65	1,402	92.5	84.7	97.4	95.4	91.0	0.84
	No	242	2,445	2,687	--	--	--	--	--	--
S. Europe	Yes	137	6	143	92.4	82.5	98.0	95.8	90.8	0.83
	No	29	287	316	--	--	--	--	--	--
Other	Yes	220	9	229	93.3	91.3	95.7	96.1	90.5	0.87
European	No	21	199	220	--	--	--	--	--	--
Asian	Yes	141	19	160	89.9	81.0	94.4	88.1	90.7	0.77
	No	33	322	355	--	--	--	--	--	--
Other	Yes	90	9	99	90.2	81.1	95.4	90.9	89.8	0.78
	No	21	185	206	--	--	--	--	--	--
Total		10,178	14,277	24,455	--	--	--	--	--	--

Statistical differences exist:
(a) for sensitivity, between "Other European" and "Aboriginal," $p<0.05$
(b) for PPV, between "Asian" and "French," $p<0.05$
(c) for PPV, between "Asian" and "British," $p<0.05$.

Greendale et al., who compared self-reported ever use of estrogen replacement therapy in a baseline questionnaire with the results of one conducted three years later, found that inaccurate

reporting was more common among non-Whites than Whites (27). In contrast, Goodman et al. found that the reporting of ever use of menopausal estrogens was more accurate among Japanese than Whites (24).

3.3.6 Language spoken

In the NPHS sample there were very few women who did not speak English, thus the reliability of the results for these women is low (Table 25). Although women who spoke English only tended to score higher on the various validity measures than women who spoke English as well as another language, none of the differences were statistically significant.

Table 25. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by language spoken

		DPIN								
Language	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
English	Yes	189	44	233	96.0	82.9	97.7	81.1	97.9	0.80
only	No	39	1,825	1,864	--	--	--	--	--	--
English &	Yes	86	27	113	95.7	80.4	97.3	76.1	97.9	0.76
other	No	21	973	994	--	--	--	--	--	--
No English	Yes	1	0	1	100.0	100.0	100.0	100.0	100.0	1.00
	No	0	19	19	--	--	--	--	--	--
Total		336	2,888	3,224	--	--	--	--	--	--

No statistical differences exist.

3.3.7 Educational attainment

Among the first time MBSP screeners, the differences in the validity measures by levels of educational attainment were small and for the most part they were not statistically significant (Table 26). While women with less than high school graduation had the lowest sensitivity, PPV and kappa, women with a degree or diploma had the lowest concordance, specificity and NPV. On the other hand, women with at least high school graduation, but not a degree or diploma, had the highest sensitivity.

Although the validity measures for women in the NPHS were also inconsistent across education levels, they were in keeping with those for the MBSP incident screeners, in that those women with the lowest education had the lowest sensitivity, PPV and kappa, but the highest concordance, specificity and NPV (Table 27). However, none of the differences in the validity measures were statistically significant.

Table 26. Self-reported current HRT use in the MBSP first time screeners compared to DPIN prescriptions, by education

		DPIN								
Education	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Less than	Yes	4,484	1,171	5,655	94.0	92.9	94.2	79.3	98.2	0.82
HS grad	No	342	19,041	19,383	--	--	--	--	--	--
HS grad	Yes	5,442	1,081	6,523	93.7	95.3	93.0	83.4	98.2	0.85
	No	266	14,412	14,678	--	--	--	--	--	--
Degree or	Yes	3,340	652	3,992	93.1	95.2	92.3	83.7	97.9	0.84
Diploma	No	168	7,788	7,956	--	--	--	--	--	--
Total		14,042	44,145	58,187	--	--	--	--	--	--

Statistical differences exist between:
(a) "Less than High School graduation" and "High School graduate" for sensitivity ($p<0.0001$), specificity ($p<0.0001$) and PPV ($p<0.0001$)
(b) "Less than High School graduation" and "Degree or Diploma" for concordance ($p<0.05$), sensitivity ($p<0.0001$), specificity ($p<0.0001$) and PPV ($p<0.0001$).

Table 27. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by education

		DPIN								
Education	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Less than	Yes	80	24	104	96.5	76.9	98.1	76.9	98.1	0.75
HS grad	No	24	1,258	1,282	--	--	--	--	--	--
HS grad &	Yes	196	46	242	95.5	84.5	97.1	81.0	97.7	0.80
Higher	No	36	1,533	1,569	--	--	--	--	--	--
Total		336	2,861	3,197	--	--	--	--	--	--

No statistical differences exist.

Table 28. Self-reported HRT use between screens in the MBSP compared to DPIN prescriptions, by education

		DPIN								
Education	MBSP	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Less than	Yes	3,085	190	3,275	91.8	82.3	97.2	94.2	90.7	0.82
HS grad	No	664	6,506	7,170	--	--	--	--	--	--
HS grad	Yes	3,814	163	3,977	92.6	86.9	97.1	95.9	90.3	0.85
	No	575	5,356	5,931	--	--	--	--	--	--
Degree or	Yes	2,083	62	2,145	93.4	88.9	97.5	97.1	90.4	0.87
Diploma	No	261	2,448	2,709	--	--	--	--	--	--
Total		10,482	14,725	24,207	--	--	--	--	--	--

Statistical differences exist:
(a) for concordance, between "Less than High School graduation" and "Degree or Diploma," $p<0.01$
(b) for sensitivity, between "Less than High School graduation" and "High school graduation," $p<0.0001$
(c) for sensitivity, between "Less than High School graduation" and "Degree or Diploma," $p<0.0001$
(d) for PPV, between "Less than High School graduation" and "High school graduation," $p<0.01$
(e) for PPV, between "Less than High School graduation" and "Degree or Diploma," $p<0.0001$.

The results for the prevalent MBSP screeners (Table 28) were more consistent than those for the incident screeners and those from the NPHS, with women who had a degree or diploma having higher scores on all validity measures except on the NPV.

Merlo et al. found that women with nine or more years of education were less accurate in reporting current HRT use than women with less education (29). A similar finding was reported by Goodman et al. for ever use of menopausal estrogens (24). In contrast, when it came to the accuracy of recall of the name of the postmenopausal estrogen that was used, West et al. found it to be much lower among women who had no college education compared to those who did (28). In a study by Greendale et al., level of education did not impact on the reliability of reporting ever use of estrogen replacement therapy (27).

3.3.8 Occupation

Although the differences were not significant, among women with an occupation, those in farming, fishing or forestry had the lowest sensitivity, NPV and kappa, but the highest specificity and PPV (Table 29). Women in processing, construction or crafts occupations had the lowest concordance, specificity and PPV and the highest sensitivity and NPV, while those in clerical, sales and service occupations had the highest concordance and kappa.

Table 29. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by occupation

		DPIN								
Occupation	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
Admin[1]	Yes	69	21	90	94.4	88.5	95.4	76.7	98.0	0.79
	No	9	437	446	--	--	--	--	--	--
Clerical[2]	Yes	79	15	94	96.2	89.8	97.2	84.0	98.3	0.85
	No	9	524	533	--	--	--	--	--	--
Farming[3]	Yes	6	0	6	92.9	60.0	100.0	100.0	92.0	0.71
	No	4	46	50	--	--	--	--	--	--
Processing[4]	Yes	10	5	15	92.4	90.9	92.7	66.7	98.4	0.73
	No	1	63	64	--	--	--	--	--	--
Not	Yes	109	30	139	96.5	75.2	98.3	78.4	98.0	0.75
applicable	No	36	1,729	1,765	--	--	--	--	--	--
Total		332	2,870	3,202	--	--	--	--	--	--

1.Administration/science/arts/religion
2. Clerical/sales/services
3. Farming/fishing/forestry
4. Processing/construction/crafts
Statistical differences exist for specificity between “Admin/Science/Arts” and “Not applicable” ($p<0.05$).

3.3.9 Income

Women with lower income had lower sensitivity, positive predictive value and kappa, but higher concordance, specificity and negative predictive value than higher income women (Table 30). However, only the differences in the PPV measure were statistically significant. Greendale et al.,

found that the reliability of reporting ever use of estrogen replacement therapy did not vary significantly by income level (27).

Table 30. Self-reported current HRT use in the NPHS compared to DPIN prescriptions, by income

		DPIN								
Income	NPHS	Yes	No	Total	Concord	Sens	Spec	PPV	NPV	Kappa
No income,	Yes	78	34	112	96.1	77.2	97.5	69.6	98.3	0.71
<$30,000	No	23	1,328	1,351	--	--	--	--	--	--
>= $30,000	Yes	152	29	181	95.5	85.4	97.2	84.0	97.5	0.82
	No	26	1,023	1,049	--	--	--	--	--	--
Not stated	Yes	46	8	54	96.4	80.7	98.3	85.2	97.7	0.81
	No	11	467	478	--	--	--	--	--	--
Total		336	2,889	3,225	--	--	--	--	--	--

Statistical differences exist between "No income, <$30,000" and ">= $30,000" for PPV ($p<0.05$).

4. Discussion / Conclusion

A substantial proportion of women attending the MBSP were current users of HRT (27.7% of women aged 50-69). Since HRT use has been reported to increase breast density and in turn reduce the sensitivity of mammographic screening (18-20), a large number of women in Manitoba may be at risk of having a misdiagnosis. Knowledge of current use of HRT may assist radiologists who are interpreting the film in making their diagnosis. In Canada, breast screening programs rely on self-report to determine HRT use.

The proportion of women aged 50-69 in the MBSP who reported they were currently using HRT (27.7%) was substantially higher than women in the NPHS (17.5%), although this is consistent with self-report in other screening programs. This difference persisted even among women who were confirmed to be current HRT users from the DPIN data. Women who attend screening programs may be more concerned about their health than the average woman. As such, they may see physicians more often and thus have a greater opportunity of being prescribed HRT.

This study has examined the validity of self-reported current use of hormone replacement therapy by comparing self-reported information in the NPHS and MBSP with data from Manitoba Health's DPIN files. Unlike questionnaires, pharmaceutical files are not affected by a respondent's recall and it has also been recognized that they provide more complete information on drug use than physician or hospital records (23,24,41). By having a population-based 'gold standard' it was possible to measure validity more accurately than most of the previous studies that have examined this issue. To our knowledge no study has examined the validity of self-reported current HRT utilization using a pharmaceutical file.

The results of this study indicate that incident and prevalent screeners in the MBSP and women in the NPHS generally report use of hormone replacement therapy accurately, although there were some differences between the three groups. For all three groups of women, concordance, specificity and the NPV were very good (90% or higher) and the kappa statistic indicated 'almost perfect' agreement. The PPV was also over 90% for MBSP prevalent screeners, but lower among

MBSP incident screeners and NPHS women. For screening programs, the sensitivity of self-report is the most important indicator of validity as it indicates the percentage of women currently using HRT who accurately report their use. Sensitivity was high among incident screeners (94.4%), but substantially lower among MBSP prevalent screeners (85.7%) and NPHS women (82.1%). Even when the NPHS sample was restricted to women aged 50-69 years the sensitivity remained low (86.5%). Part of the variations in sensitivity may have arisen from the differences in the questions about HRT use. In the MBSP incident screeners were asked about current use, while prevalent screeners were asked about use since their last screen. NPHS women were asked about use in the last month.

Although there were some differences, the above patterns tended to be fairly consistent across the various population subgroups examined (e.g sensitivity was generally highest among incident screeners and lowest among NPHS women regardless of the subgroup). Furthermore, for sensitivity, the same population groups tended to score lower in all three samples of women. Women who were elderly, lived in rural areas, or who had less than a high school education had the lowest sensitivity among MBSP incident and prevalent screeners and among the NPHS sample. Aboriginal women in the MBSP also had lower sensitivity than non-Aboriginal women. The reverse was found in the NPHS, but that result was based on a small sample. In this study only univariate analyses of validity were undertaken. Future research should take advantage of the large number of women in the NPHS and MBSP sample and undertake multivariate analyses to identify the profile of women who do not accurately self-report their HRT use.

Although the DPIN database was considered to be a ‘gold standard’, it does have some limitations. There may be some women who are in the DPIN, but who correctly stated they were not current users of HRT. The DPIN data identifies prescriptions that were filled, however, there is no guarantee that the medication was taken. Although West et al. found that very few respondents who were dispensed a drug did not use it (26). Conversely, there may be women who are actually using some form of HRT but who were not recorded in the DPIN database. For example, it is possible to obtain plant-based hormones (e.g. progesterone yam cream) without a prescription, and some women may also have obtained a free sample of drugs from their physician. Also included in this group would be women who may be using HRT but are not eligible for inclusion in DPIN, such as Status Aboriginals. A small number of women may also have received a prescription out-of-province.

References

1. Banks E. Hormone replacement therapy and the sensitivity and specificity of breast cancer screening: a review. J Med Screen 2001;8:29-35.
2. Lane JM, Russell L, Khan SN. Osteoporosis. Clin Orthop 2000;372:139-50.
3. Barrett Connor E, Grady D. Hormone replacement therapy, heart disease and other considerations. Ann Rev Public Health 1998;19:55-72.
4. Yaffe K, Sawaya G, Lieberburg I, Grady D. Estrogen therapy in postmenopausal women: effects on cognitive function and dementia. JAMA 1998;279:688-95.
5. Shaywitz BA, Shaywitz SE. Estrogen and Alzheimer disease. JAMA 2000;283:1055-6.
6. Statistics Canada. National Population Health Survey 1994-95 Public Use Microdata File.
7. Menashe Y, Judd HL. Estrogen-progestin replacement regimens: a critical analysis. In: Adashi EY, Rock JA, Resenwaks Z, eds. Reproductive Endocrinology, Surgery, and Technology. Philadelphia: Lippincott-Raven, 1995:1759-73.
8. Jain MG, Rohan TE, Howe GR. Hormone replacement therapy and endometrial cancer in Ontario, Canada. J Clinic Epidemiol 2000;53:385-91.
9. Cancer CGoHFiB. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52 705 women with breast cancer and 108 411 women without breast cancer. Lancet 1997;350:1047-59.
10. Persson I, Thurfjell E, Bergstrom R, Holmberg L. Hormone replacement therapy and the risk of breast cancer. Nested case-control study in a cohort of Swedish women attending mammography screening. Int J Cancer 1997;72:758-61.
11. Persson I, Weiderpass E, Bergkvist L, Bergstrom R, Schairer C. Risks of breast and endometrial cancer after estrogen and estrogen-progestin replacement. Cancer Causes Control 1999;10:253-60.
12. Colditz GA, Rosner B. Use of estrogen plus progestin is associated with greater increase in breast cancer risk that estrogen alone. Am J Epidemiol 1998;147:64S.
13. Schairer C, Lubin J, Troisi R, Sturgeon S, Brinton L, Hoover R. Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer. JAMA 2000;283:485-491.
14. Laya MB, Larson EB, Taplin SH, White E. Effect of estrogen replacement therapy on the specificity and sensitivity of screening mammography. J Natl Cancer Inst 1996;88:643-9.
15. Harvey JA, Pinkerton JV, Herman CR. Short-term cessation of hormone replacement therapy and improvement of mammographic specificity. J Natl Cancer Inst 1997;89:1623-5.
16. Kavanaugh AM, Mitchell H, Giles GG. Hormone replacement therapy and accuracy of mammographic screeening. Lancet 2000;355:270-4.
17. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081-7.
18. Stomper PC, Van Voorhis BJ, Ravnikar VA, Meyer JE. Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study. Radiology 1990;174:487-90.
19. Laya MB, Gallagher JC, Schreiman JS, Larson EB, Watson P, Weinstein L. Effect of postmenopausla hormone replacement therapy on mammographic density and parenchymal pattern. Radiology 1995;196:433-7.

20. Cohen MEL. Effect of hormone replacemetn therapy on cancer detection by mammography. Lancet 1997;349:1624.
21. Horowitz RI, Feinstein AR, Stremlau JR. Alternative data sources and discrepant results in case-control studies of estrogens and endometrial cancer. Am J Epidemiol 1980;111:389-94.
22. Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. Am J Epidemiol 1982;116:114-22.
23. Persson I, Bergkvist L, Adami H-O. Reliability of women's histories of climacteric oestrogen treatment assessed by prescription forms. Int J Epidemiol 1987;16: 222-8.
24. Goodman MT, Nomura AMY, Wilkens LR, Kolonel LN. Agreement between interview information and physician records on history of menopausal estrogen use. Am J Epidemiol 1990;131:815-25.
25. Jannausch ML, Sowers MFR. Consistency of perimenopausal estrogen use reporting by women in a population-based prospective study. Maturitas 1992;14:161-9.
26. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. Am J Epidemiol 1995;142:1103-12.
27. Greendale GA, James MK, Espeland MA, Barrett-Connor E. Can we measure prior postmenopausal estrogen/progestin use? The postmenopausal estrogen/progestin interventions trial. Am J Epidemiol 1997;149:763-70
28. West SL, Savitz DA, Koch G, et al. Demographics, health behaviors, and past drug use as predictors of recall accuracy for previous prescription medication use. J Clin Epidemiol 1997;50:975-80.
29. Merlo J, Berglund G, Wirfält E, Gullberg B, Hedblad B, Manjer J, Hovelius B, Janzon L, Hanson BS, Östergren PO. Self-administered questionnaire compared with a personal diary for assessment of current use of hormone therapy: an analysis of 16,060 women. Am J Epidemiol 2000;152:788-92.
30. Canada S. National Population Health Survey, 1996/97: Public Use Microdata Files Catalogue 82M0009XCB, CD-ROM and 82M0009GPE, documentation). Ottawa: Minister of Industry, 1998.
31. Metge C, Black C, Peterson S, Kozyrskyj A, Roos NP, Bogdanovic B. Analysis of Patterns of Pharmaceutical Use in Manitoba, 1996: Key Findings. Winnipeg: The Manitoba Centre for Health Policy and Evaluation, December 1999.
32. The Manitoba Centre for Health Policy and Evaluation http://www.umanitoba.ca/centres/mchpe/1mchpe.htm .
33. Brownell M, Martens P, Kozyrskyj A, Fergusson P, Lerfald J, Mayer TM, Derksen SA, Friesen D. Assessing the Health of Children in Manitoba: A Population-Based Study. Winnipeg:The Manitoba Centre for Health Policy and Evaluation, March 2001.
34. Canadian Pharmaceutical Association. CPS, Compendium of Pharmaceuticals and Specialities. Ottawa, Ontario, Canada: Bruce LD, Canadian Pharmaceutical Association, 2000.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
36. Miller JJ. The inverse of the Freeman-Tukey double arcsine transformation. Amer Stat 1978;32:138.

37. Westfall PH, Young SS. P-value adjustments for multiple tests in multivariate binomial models. J Amer Stat Assoc 1989;84:780-6.
38. Westfall PH, Young SS. Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustments. New York: John Wiley & Sons, 1993.
39. The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. Breast Cancer Res 1999;1:73-80.
40. Van den Brandt PA, Petri H, Dorant E, Goldbohm A, Van de Crommert S. Comparison of questionnaire information and pharmacy data on drug use. Pharm Weekbl [Sci] 1991;13:91-6.
41. West SL, Strom BL, Freundlich B, Normand E, Koch G, Savitz DA. Completeness of prescription recording in outpatient medical records from a health maintenance organization. J Clin Epidemiol 1994;47:165-71.