ORIGINAL RESEARCH

Hydatidiform mole: validity of the registration in the Danish National Patient Registry, the Danish Cancer Registry, and the Danish Pathology Registry 1999–2009

Helle Lund[1,2]
Mogens Vyberg[1,2]
Helle Højmark Eriksen[3]
Anni Grove[1]
Annette Østergaard Jensen[1]
Lone Sunde[4,5]

[1]Institute of Pathology, Aalborg University Hospital, Aalborg, Denmark; [2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; [3]Unit of Epidemiology and Biostatistics, Aalborg University Hospital, Aalborg, Denmark; [4]Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark; [5]Department of Biomedicine, Aarhus University, Aarhus, Denmark

Correspondence: Helle Lund
Institute of Pathology, Aalborg University Hospital, Ladegaardsgade 3, 9000 Aalborg, Denmark
Tel +45 9766 5284
Fax +45 9766 5320
Email helu@rn.dk

Purpose: To examine the validity of registration of hydatidiform mole (HM) in the Danish National Patient Registry (NPR), the Danish Cancer Registry (DCR), and the Danish Pathology Registry (DPR).

Patients and methods: We selected women registered with a first-time HM code in NPR, DCR, and DPR from 1999 to 2009. We found most women registered in DPR. For a random sample of women registered in DPR, the coding was validated by comparing with the pathology report. Completeness and positive predictive value (PPV) of registration with an HM code in NPR and DCR were calculated using DPR as the reference. Details of women registered in NPR or DCR, but not in DPR, were scrutinized.

Results: In NPR and DPR, 1,520 women were identified in total; 1,057 (70%) were found in both registries, 65 (4%) only in NPR, and 398 (26%) only in DPR. In DCR and DPR, 1,498 women were identified in total; 1,174 (78%) in both registries, 47 (3%) only in DCR, and 277 (19%) only in DPR. For 149/150 randomly selected women registered with an HM code in DPR (99%), the pathology report was consistent with the diagnosis of HM. Completeness of NPR was 73% (95% CI: 70%–75%) and PPV was 94% (95% CI: 93%–95%). Completeness of DCR was 72% (95% CI: 69%–75%) in 1999–2003 and 90% (95% CI: 87%–92%) in 2004–2009. PPV of DCR was 96% (95% CI: 95%–97%) throughout the period.

Conclusion: Validation of registry data is important before using these. For research on the number of HMs in Denmark, DPR is the most valid data source. NPR and DCR appear to be equally valid before 2004. However, for research after 2004, DCR should be preferred rather than NPR.

Keywords: epidemiology, completeness, positive predictive value, complete hydatidiform mole, partial hydatidiform mole, agreement

Introduction

Hydatidiform mole (HM) is an abnormal pregnancy and the most common form of gestational trophoblastic disease. HMs are characterized by vesicular swelling of the chorionic villi, hyperplasia of the trophoblastic layer, and absence or abnormal development of the fetus. HMs can be classified into two types, complete hydatidiform mole (CHM) and partial hydatidiform mole (PHM).[1] The incidences of HM reported in different parts of the world are highly varying. In Europe and North America, the reported HM incidences range from 66 to 121 per 100,000 pregnancies, while much higher incidences have been reported in Asia and the Middle East.[2] The causes of the

1223

varying incidences reported have not been established, but it has been suggested that diet and nutrition may contribute to the etiology.[3] Comparing incidence rates of HM from different studies is difficult. The population at risk may be inadequately described, the methodological design varies, there may be differences in disease definition and changes in diagnostic methods over time.[4–6] Previous studies have demonstrated underreporting of HMs to the Swedish Cancer Register with only very little improvement over time.[7,8] In Denmark, women with an HM are registered in three governmental registries: the Danish National Patient Registry (NPR),[9] the Danish Cancer Registry (DCR),[10] and the Danish Pathology Registry (DPR).[11] The aim of this study was to examine the validity of the registration of HMs in the NPR, the DCR, and the DPR from 1999 to 2009.

Patients and methods
Registries
The NPR is an electronic medical registry storing data of all patients discharged from Danish somatic hospitals nationwide since 1977.[12] Data on outpatients have been registered since 1995. Standardized data from regional Patient Administrative Systems are submitted electronically to the NPR at least monthly, but in practice often weekly or daily.[13] The information reported to the NPR includes administrative data and data on diagnoses, examinations, and treatments (http://www.esundhed.dk/dokumentation/Registre/Sider/Register.aspx). The 10th revision of the International Classification of Diseases (ICD10) has been used to classify diseases since 1994.[9]

The DCR has recorded primary cases of cancer and related diseases (including HMs) on a nationwide basis since 1943, and reporting has been mandatory since 1987. The information recorded in the DCR includes both personal and tumor characteristics (http://www.esundhed.dk/dokumentation/Registre/Sider/Register.aspx). In 2004, the DCR went through a process of modernization, which included electronic notifications from the NPR and an automated cancer logic algorithm based on ICD10 replacing former paper notifications from hospitals. Furthermore, data registered in the period 1978–2004 were re-coded using ICD10.[14] Approximately, 80%–90% of the manual coding has been replaced by the automatic cancer logic.[10] Reminders are regularly dispatched to the hospitals who fail to report.

Computer-based recording of pathology data began in Denmark in the 1970s. The DPR was established in 1997. Since 1999, reporting in the DPR automatically takes place through the Danish Pathology Data Bank (DPB), a nationwide database, working as a routine online tool for all pathology departments in Denmark. For each specimen, patient data, date of registration, the requesting and investigating hospitals and departments, gross and microscopy descriptions, and diagnoses are registered. Coding is performed using the Danish modification of the Systemized Nomenclature of Medicine (SNOMED, http://www.patobank.dk/). The data quality is ensured by the approval of all diagnostic statements in the pathology report by a pathologist and a debugging system, which ensures that all statements are given at least one code for topography and one code for morphology. When a new pathology report is completed in the DPB, data are automatically sent to the DPR, which is updated on a daily basis.[11,15]

Using the 10-digit civil personal registration number[16,17] assigned to all Danish citizens at birth or immigration since 1968, unambiguous data linkage between the abovementioned registries is possible.

Identification of women
We initially retrieved data on all women registered with an HM code in the NPR, the DCR, and/or the DPR from 1.1.1987 to 31.12.2010. In the NPR, the DCR, and the DPR, women were registered with their civil personal registration number, the date of registration, and an HM code, along with a number of other variables. It is possible for a woman to be registered with an HM code more than once in the NPR and the DPR, but a woman with an HM is registered only with her first HM in the DCR. To make the data sets comparable, we adjusted the data sets from the NPR and the DPR: a woman registered with an HM code more than once in a registry was included in the data set from that registry with the date of the first HM code only.

To validate the registration of women with an HM code in the NPR, the DCR, and the DPR in the period 1999–2009, we selected data on women registered with an HM code for the first time in the above data sets from 1.1.1999 to 31.12.2009. The codes used for identification of the women with an HM in the various registries are listed in Table 1. In the DPR, code M91000 most likely was intended to indicate a woman diagnosed with a CHM. However, as no other code was available, M91000 may also have been used to indicate that the women had an "HM, not otherwise specified". In the Danish SNOMED, the code ÆYYY00 is a moderator indicating "suspicion of". The code can be used as a supplement to the morphology codes.

A woman was classified as being registered identically in two registries, if she was registered in both registries and the difference between the two dates of registration in the two

Table 1 Codes for HM in the DPR, the DCR, and the NPR in 1999–2009[a]

DPR (SNOMED)		DCR (ICD10)		NPR (ICD10)	
Code	**Diagnosis**	**Code**	**Diagnosis**	**Code**	**Diagnosis**
M91000	Hydatidiform mole	DO010	Complete hydatidiform mole	DO010	Complete hydatidiform mole
		DO019	Hydatidiform mole, not otherwise specified	DO019	Hydatidiform mole, not otherwise specified
M91030 ÆYYY00[b]	Partial hydatidiform mole A moderator indicating "suspicion of"	DO011	Partial hydatidiform mole	DO011	Partial hydatidiform mole
				DO019A	Intrauterine hydatidiform mole
				DO020B	Extrauterine hydatidiform mole

Notes: [a]The table also illustrates the corresponding HM codes for calculating agreement and kappa values for registration of morphologic subdiagnoses of HMs. [b]The moderator code ÆYYY00 can be used as a supplement to the morphology code (M91000 or M91030).
Abbreviations: DCR, Danish Cancer Registry; DPR, Danish Pathology Registry; HM, hydatidiform mole; ICD10, 10th revision of the International Classification of Diseases; NPR, National Patient Registry; SNOMED, Systemized Nomenclature of Medicine.

registries was <6 months. Two and one women registered in 1998 with an HM code in the NPR and the DCR, respectively, were included in the data sets from these registries, as these women were registered in the DPR in 1999, and there was <6 months between the two dates of registration. Likewise, one woman registered in both the NPR and the DCR in 2010 was included in the data sets from these registries, as she was registered in the DPR in 2009 <6 months earlier (Figure 1).

Including women with their first HM code only could introduce errors. For instance, if a woman actually had had two HMs, and both HMs were coded as HMs in the DPR, whereas only the second HM was coded as an HM in the NPR, we would underestimate both completeness and positive predictive value (PPV) of the NPR relative to the DPR. To compensate for this, we next excluded women from the data sets from the NPR and the DPR, and the DCR and the DPR who had >6 months between the dates of registration in the two registries. This led to exclusion of three women from the data set from the NPR and the corresponding data set from the DPR, and seven women from the data set from the DCR and the corresponding data set from the DPR (Figure 1).

Women present in the data set from the NPR and/or the DCR, but not present in the data set from the DPR, were researched in the DPR 4 years after the primary search, and the causes explaining the difference were categorized. Women present in the data set from the DPR, but not present in the data set from the NPR, were researched in the original data set from the NPR, to identify women excluded during the reduction of the data sets to include women with first-time HM codes only.

To evaluate the validity of the HM coding in the DPR, we checked the pathology report for 150 randomly selected women with an HM according to the SNOMED coding in the period from 1999 to 2009. The pathology reports were reviewed by two pathologists, with special interest in gynecological pathology (HL and AG). Based on the description in the pathology report, a conceptus was classified as an HM if the conclusion stated that the conceptus was an HM and/or if the report listed morphological findings diagnostic for an HM.[18] Next, the conceptuses identified as HMs were categorized into four groups: CHM, PHM, HM-NOS and HM, subtype not specified. HMs for which the conclusion in the pathology report stated a CHM or PHM were classified as such. An HM for which the report stated that subtyping had not been possible, was classified as an HM-NOS. If there was no clear conclusion regarding the subtype of HM in the pathology report, the reviewing pathologists classified the subtype if possible using the histopathological features of HMs described by Sebire.[18] The fourth group represents HMs, for which the text was not sufficiently informative for the reviewing pathologists to perform subtyping (HM, subtype not specified).

Statistical analysis

As no pathology report can be signed out without registration in the DPR, the proportion of missing data in this registry is expected to be low.[11] Concordantly, comparing registration in the NPR, the DCR, and the DPR, more women were registered in the DPR than in the NPR and the DCR (details are provided in the "Results" section). Although more accurate sonographic diagnoses, especially for CHMs, have been observed,[19] histopathological examination of conceptuses remains the current gold standard for the identification of gestational trophoblastic diseases.[20] Furthermore, 149/150 women registered with an HM code in the DPR had an HM according to the pathology report (details are provided in the "Results" section). Therefore, we used data from the DPR as the reference standard calculating the completeness and PPV of data from the NPR and the DCR. The completeness
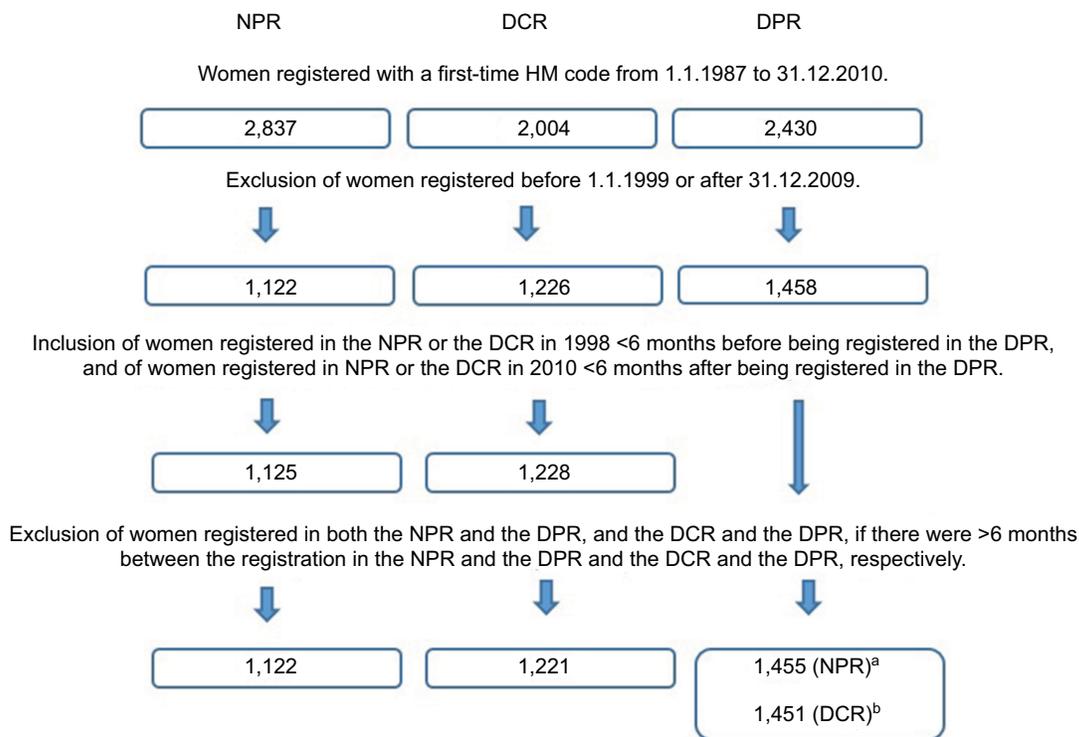
NPR	DCR	DPR

Women registered with a first-time HM code from 1.1.1987 to 31.12.2010.

2,837	2,004	2,430

Exclusion of women registered before 1.1.1999 or after 31.12.2009.

1,122	1,226	1,458

Inclusion of women registered in the NPR or the DCR in 1998 <6 months before being registered in the DPR, and of women registered in NPR or the DCR in 2010 <6 months after being registered in the DPR.

1,125	1,228	

Exclusion of women registered in both the NPR and the DPR, and the DCR and the DPR, if there were >6 months between the registration in the NPR and the DPR and the DCR and the DPR, respectively.

1,122	1,221	1,455 (NPR)[a] 1,451 (DCR)[b]

Figure 1 Data retrieval of women registered with an HM code in the NPR, the DCR, and the DPR. Two different data sets from DPR were created. [a]For comparing with the data set from the NPR, three women registered in the NPR and the DPR with >6 months between the dates of registration were excluded. [b]For comparing with the DCR, seven women registered in the DCR and the DPR with >6 months between the dates of registration were excluded.
Abbreviations: DCR, Danish Cancer Registry; DPR, Danish Pathology Registry; HM, hydatidiform mole; NPR, National Patient Registry.

of the NPR and the DCR was calculated as the number of women registered with an HM code in the data set from the DPR, who were also found with an HM code in the data sets from the NPR and the DCR, respectively, divided by the total number of women registered with an HM code in the data set from the DPR. Likewise, the PPV was calculated as the proportion of women registered with an HM code in both of the data sets from the DPR and the NPR, or the DPR and the DCR, to the women registered with an HM code in the data set from the NPR or the DCR, respectively.

To determine the degree of consensus between the morphologic subdiagnoses of HMs registered in the NPR and the DPR, and the DCR and the DPR, respectively, the agreement and Cohen's kappa values were calculated.[21] The codes for morphologic subdiagnoses accepted as corresponding in the three registries are shown in Table 1. Statistical analyses were performed with STATA software version 13 (StataCorp LLC. College Station, USA). The study was approved by the Danish Protections Agency (j. nr. 2014-41-3541).

Results

In the period 1.1.1999–31.12.2009, 1,520 women were registered for the first time with an HM code in the NPR and/

or registered for the first time in the DPR with an HM code. One thousand fifty-seven women (70%) were registered in both registries, 65 women (4%) were registered in the NPR only, and 398 women (26%) were registered in the DPR only (Table 2). Likewise, in the period 1.1.1999–31.12.2009, 1,498 women were registered for the first time in the DCR with an HM code and/or registered for the first time in the DPR with an HM code. One thousand one hundred seventy-four women (78%) were registered in both registries, 47 (3%) were registered only in the DCR, and 277 (19%) were registered only in the DPR (Table 3). Thus, the DPR was the registry with the highest number of women registered with an HM code.

In the DPR, the moderator code ÆYYY00 indicating "suspicion of" had been added to the morphology code for 15% of the women (Table 4). No significant difference was observed regarding the use of the moderator code in the period before and after 2004 (data not shown).

For 150 randomly selected women, 149 registered with an HM code in the DPR had an HM according to the pathology report (Table 5). For one woman, there was discrepancy between the report and the coding. The conceptus was initially described as a suspected PHM and coded M91030 + ÆYYY00. Following flow cytometry, the pathologist concluded in the pathology report that the conceptus was not an

Table 2 Number of women with a first-time hydatidiform mole code in the NPR and/or the DPR in 1999–2009

Data set from DPR[a]	Women registered in			Total, n
	Both NPR and DPR, n (%)	Only NPR, n (%)	Only DPR, n (%)	
All	1,057 (70)	65 (4)	398 (26)	1,520
Without ÆYYY00	954 (68)	168 (12)	286 (20)	1,408

Notes: [a]All: Data set including women registered with the SNOMED code M91000 (hydatidiform mole) or M91030 (partial hydatidiform mole), with or without ÆYYY00 (suspicion of). Without ÆYYY00: Data set including only women registered with the SNOMED code M91000 or M91030, without ÆYYY00.
Abbreviations: DPR, Danish Pathology Registry; NPR, National Patient Registry; SNOMED, Systemized Nomenclature of Medicine.

Table 3 Number of women with a first-time hydatidiform mole code in the DCR and/or the DPR in 1999–2009

Data set from DPR[a]	Women registered in			Total, n
	Both DCR and DPR, n (%)	Only DCR, n (%)	Only DPR, n (%)	
All	1,174 (78)	47 (3)	277 (19)	1,498
Without ÆYYY00	1,088 (79)	133 (10)	148 (11)	1,369

Notes: [a]All: Data set including women registered with the SNOMED code M91000 (hydatidiform mole) or M91030 (partial hydatidiform mole), with or without ÆYYY00 (suspicion of). Without ÆYYY00: Data set including only women registered with the SNOMED code M91000 or M91030, without ÆYYY00.
Abbreviations: DCR, Danish Cancer Registry; DPR, Danish Pathology Registry; SNOMED, Systemized Nomenclature of Medicine.

Table 4 Number of women registered with a first-time hydatidiform mole code in the Danish Pathology Registry in 1999–2009[a]

SNOMED code M91000			SNOMED code M91030			Total, n
Without ÆYYY00, n (%)	With ÆYYY00, n (%)	All, n (%)	Without ÆYYY00, n (%)	With ÆYYY00, n (%)	All, n (%)	
552 (85)	99 (15)	651 (100)	691 (86)	116 (14)	807 (100)	1,458

Notes: [a]The data set presented here is larger than the data sets presented in Tables 2 and 3, as three and seven women, respectively, were excluded from the latter data sets because the dates of registration in two registries differed >6 months (details are provided in the "Patients and methods" section). The code M91000 represents hydatidiform mole, M91030 represents partial hydatidiform mole, and ÆYYY00 represents a moderator indicating "suspicion of".
Abbreviation: SNOMED, Systemized Nomenclature of Medicine.

Table 5 Concordance between the description in the pathology report and the SNOMED code in 150 randomly selected women registered with a first-time HM code in Danish Pathology Registry in 1999–2009

Code	Information in the pathology report						
	Not mole	HM, subtype[a]			Total, subtype specified	HM, subtype not specified[b]	Total, subtype specified or not specified
		CHM	PHM	HM- NOS			
M91000	0	46	11	8	**65**	8	**73**
M91030	1	0	76	0	**77**	0	**77**
Total	**1**	**46**	**87**	**8**	**142**	**8**	**150**

Notes: [a]The HMs were categorized according to the pathology report into CHMs, PHMs, or HM-NOS. [b]HMs for which the pathology report was not sufficiently informative to the reviewing pathologists to identify the subtype.
Abbreviations: CHMs, complete hydatidiform moles; HM, hydatidiform mole; HM-NOS, HMs for which subtyping was not possible; PHMs, partial hydatidiform moles; SNOMED, Systemized Nomenclature of Medicine.

HM, but the coding remained unchanged. For 76/77 women registered with the code M91030, the pathology report stated that they had a partial hydatidiform mole. Forty-six of 73 women registered with the code M91000 had a CHM according to the pathology report.

Using the data set from the DPR as the reference standard, the completeness of the NPR from 1.1.1999 to 31.12.2009 was 73% (95% CI: 70%–75%) and the PPV was 94% (95% CI: 93%–95%; Table 6). There were no significant

differences between the completeness and PPV for the periods 1.1.1999–31.12.2003 and 1.1.2004–31.12.2009 (data not shown). When excluding women registered with an HM code with the moderator code ÆYYY00 from the data set from the DPR, the completeness of NPR improved to 77% (95% CI: 75%–79%), while the PPV lowered to 85% (95% CI: 83%–87%).

In the period 1.1.1999–31.12.2003, the completeness of the DCR relative to the DPR was 72% (95% CI: 69%–75%).

Table 6 Completeness and PPV for registration of a woman with a hydatidiform mole in the NPR and the DCR using registration in the DPR as reference standard

Registry	Period	Data from DPR[a]	Completeness (95% CI), %	PPV (95% CI), %	n[b]
NPR	1999–2009	All	73 (70–75)	94 (93–95)	1,520
		Without ÆYYY00	77 (75–79)	85 (83–87)	1,408
DCR	1999–2009	All	81 (79–83)	96 (95–97)	1,498
		Without ÆYYY00	88 (86–90)	89 (87–91)	1,369
	1999–2003	All	72 (69–75)	95 (93–97)	770
		Without ÆYYY00	79 (75–82)	90 (87–92)	702
	2004–2009	All	90 (87–92)	97 (95–98)	728
		Without ÆYYY00	98 (97–99)	89 (86–91)	667

Notes: [a]All: Data set including women registered with the SNOMED code M91000 (hydatidiform mole) or M91030 (partial hydatidiform mole), with or without ÆYYY00 (suspicion of). Without ÆYYY00: Data set including only women registered with the SNOMED code M91000 or M91030, without ÆYYY00. [b]The number of women included in the data set from the registry indicated or in the data set from the DPR or both, for the period indicated.
Abbreviations: DCR, Danish Cancer Registry; DPR, Danish Pathology Registry; NPR, National Patient Registry; PPV, positive predictive value; SNOMED, Systemized Nomenclature of Medicine.

There was an improvement in the completeness to 90% (95% CI: 87%–92%) for the period 1.1.2004–31.12.2009. The PPV was 96% (95% CI: 95%–97%), with no significant difference between the periods before and after 1.1.2004 (Table 6). When excluding women registered with HMs with the moderator code ÆYYY00 in the DPR, the completeness of the DCR improved to 98% (95% CI 97%–99%) for the period 1.1.2004–31.12.2009, and the PPV for this period lowered to 89% (95% CI: 86%–91%).

A total of 76 women were identified with an HM code in the NPR and/or the DCR but not in the DPR (Table 7). Scrutinizing the data for these, we identified 33 women for whom the data in the DPR indicated that the women had had a nonmolar conceptus in the relevant period, and nine women for whom the discrepancy could be explained by the HM in question not being the first HM of the patient, the registrations in two registers being made before and after the start of the study period, respectively, or the diagnosis being made in a non-Danish citizen. However, for 18 women, an HM code was apparently truly missing in our first data set: 11 women apparently had been diagnosed with an HM in the period, however, this was only registered after we retrieved the data set and 7 women had been diagnosed with an HM according to the pathology report, but an unspecific (n=6) or incorrect (n=1) code had been chosen. For 16 women, no relevant data were available in the DPR. Based on these observations, the completeness of the DPR would be between 1,458/(1,458+18)=98.8% and 1,458/(1,458+34)=97.7%. No women were "missing" in the NPR data set due to HMs diagnosed before the study period (data not shown).

Of the 1,057 women registered for the first time with an HM code in both the NPR and the DPR, 1,046 women were registered with codes indicating morphologic subtypes of HM that could be related (11 women were excluded, as they were registered in the NPR with the ICD10 codes DO019A or DO020B, which could not be related to the codes of HM in the DPR; Table 1). The agreement between the subtypes registered in the NPR and the DPR was 84% (95% CI: 82%–86%), and the kappa value was 0.68 (95% CI: 0.64–0.73) with no significant difference between the periods 1.1.1999–31.12.2003 and 1.1.2004–31.12.2009 (Table 8). For the 1,174 women registered for the first time with an HM code in both the DCR and the DPR, the agreement between the morphologic subtypes registered increased from 71% (95% CI: 67%–75%) in the period 1.1.1999–31.12.2003 to 77% (95% CI: 73%–80%) in 1.1.2004–31.12.2009 (P=0.02). In the same period, the kappa value increased from 0.45 (95% CI: 0.38–0.51) to 0.54 (95% CI: 0.47–0.60; Table 9).

Discussion

In this study, we examined the validity of registration of women with HMs in three national Danish registries, the NPR, the DCR, and the DPR.

The strength of our study is that the health care system in Denmark ensures equal access to free health care to all residents, and clinicians in the public health care system are obliged to register patients in a number of nationwide databases. Thus, we were able to compare the registration of women with HMs in not only two, but in three nation-wide health registries. Furthermore, we were able to link data using the Danish civil personal registration number that uniquely identifies every Danish citizen.

A limitation to our study is that we reduced the data sets from the NPR and the DPR to women with a first-time HM code, and that we further excluded women registered for the first time with an HM code in the NPR and the DPR, or in

Table 7 Characterization of women registered with an HM code in the NPR (n=65) and the DCR (n=47), but not registered with an HM code in the DPR

Characteristics	Number of women identified in the registry indicated, but not in the DPR			
	Both NPR and DCR	Only NPR	Only DCR	Total
The woman was not registered with a conception in the DPR in the relevant period	10	2	4	16
The woman was not registered with an HM code in the DPR when the original data set was retrieved, but was identified when the DPR was researched[a]	4	4	3	11
The woman was registered with a conception in the DPR, but the morphological findings were not diagnostic for HM according to the pathology report	14	16	3	33[b]
The woman was registered in the DPR with a code indicating a pregnancy, but not specifying that this was an HM, whereas the description in the pathology report was consistent with a diagnosis of HM	7[c]	0	0	7
Other causes	1[d]	7[e]	1[f]	9
Total	36	29	11	76

Notes: [a]The DPR was researched 4 years after the original search. [b]For 6/33 women, there was some discussion regarding trophoblastic diseases in the pathology report. [c]For 6/7 women, the coding was made by the same pathologist, who used the code for "abnormal pregnancy product" even though the pathologist used the term "hydatidiform mole" in the description in the pathology report. [d]The woman was diagnosed with an HM outside Denmark and apparently subsequently treated and registered in both the NPR and the DCR in Denmark. [e]One woman was diagnosed with an HM within the relevant period, but was excluded from our DPR data set as the patient had a previous mole; two women were registered in the DPR in late December 1998 and therefore not included in our DPR data set, but registered in the NPR between January and February 1999; four women were registered with a temporary CPR number in the NPR, but they were not registered in the DPR. [f]The woman was diagnosed with an HM within the relevant period, but was excluded from our DPR data set as the patient had a previous mole.
Abbreviations: DCR, Danish Cancer Registry; DPR, Danish Pathology Registry; HM, hydatidiform mole; NPR, National Patient Registry.

Table 8 Agreement and kappa values for morphologic subdiagnoses of hydatidiform mole registered in the Danish National Patient Registry and the Danish Pathology Registry

Period[a]	Agreement	95% CI	Kappa	95% CI
1999–2009	84%	82%–86%	0.68	0.64–0.73
1999–2003	83%	80%–86%	0.66	0.60–0.72
2004–2009	85%	82%–88%	0.71	0.65–0.77

Notes: [a]Period: 1.1.1999–31.12.2009, 1.1.1999–31.12.2003, and 1.1.2004–31.12.2009.

Table 9 Agreement and kappa values for morphologic subdiagnoses of hydatidiform mole registered in the Danish Cancer Registry and the Danish Pathology Registry

Period[a]	Agreement	95% CI	Kappa	95% CI
1999–2009	74%	72%–77%	0.50	0.45–0.54
1999–2003	71%	67%–75%	0.45	0.38–0.51
2004–2009	77%	73%–80%[b]	0.54	0.47–0.60

Notes: [a]Period: 1.1.1999–31.12.2009, 1.1.1999–31.12.2003, and 1.1.2004–31.12.2009. [b]P=0.02.

the DCR and the DPR on dates that were >6 months apart, to reduce errors caused by the reduction of the data sets to women with a first-time HM code. However, this affected only three and seven women, respectively.

We found indications that the DPR was the most valid register. The DPR had most women registered with an HM code, and the concordance between the description in the pathology report and the codes in the DPR was very good.

This was not surprising because HM is a morphological diagnosis that in principle should be made by the pathologist, and registration in the DPR is linked to signing out of the pathology report. Nevertheless, we disclosed some women registered with an HM in the NPR or the DCR, but not in the DPR. Furthermore, we cannot exclude that some women were diagnosed with an HM and not registered with an HM code in any of the three registries. However, a substantial fraction of the women registered with an HM code in the NPR and/or the DCR but not in the DPR were registered in the DPR in the relevant period with a non-HM code, suggesting that these women were erroneously registered with an HM code in the NPR or the DCR. Furthermore, as almost 99% of the women registered with an HM code in the NPR and/or the DPR, and almost 99% of the women registered with an HM code in the DCR and/or the DPR were actually registered with an HM code in the DPR, the frequency of missing registrations in the DPR is likely very small.

The finding that almost half of the women "missing" in the DPR actually were diagnosed and registered in the DPR with morphological findings in a conceptus not diagnostic for HM may be explained by coding in the NPR and the DCR before receiving the final pathology report. The miscoding caused by one pathologist repeatedly using the less-specific code "abnormal pregnancy product", even though the text in the pathology report was consistent with a diagnosis of HM, illustrates the importance of using national coding guidelines.

In the period 1999–2003, both the NPR and the DCR demonstrated a completeness of 72% compared with the DPR. A comparable underreporting of 20%–25% in the Swedish Cancer registry with little improvement over time, has been documented for the periods 1971–1986[8] and 1973–2004,[7] although reporting to the cancer registry was mandatory in Sweden. As molar diseases are relatively rare and diagnosed in both specialized and in nonspecialized hospitals, the observed underreporting may be explained by lack of knowledge that HMs should be registered in the same registry as malignant diseases. Also, doctors at nonspecialized hospitals who see only few cases of hydatidiform mole during their professional life, may have a tendency to choose a wrong or less-specific code. The marked improvement in completeness and agreement for morphological subdiagnoses in the DCR compared with the DPR after 2003, corresponds well with the improved method of registration in the DCR that is used since 2004.

The agreement and kappa values for registration of morphologic subdiagnoses of HM in the NPR relative to the DPR were higher than the agreement and kappa values for the DCR relative to the DPR throughout the period. This was unexpected, as since 2004 most of the data in the DCR originate from the NPR. However, the higher agreement for the NPR may be related to the lower completeness of the NPR – possibly a lower proportion of women diagnosed with diagnostically difficult cases of HM are registered in the NPR.

As PHMs can be difficult to separate morphologically and immunohistochemically from hydropic abortions,[22,23] one could have expected that the moderator ÆYYY00 indicating some uncertainty regarding the diagnosis of HM, had been used more often for PHMs compared with CHMs. An increased use over time could have also been expected, due to the evacuation at earlier gestational ages with the advent of more sensitive ultrasound scanners over time. When an HM is terminated early, both the sonographic findings[20] and the histopathological findings are more discrete,[24,25] causing difficulties in diagnosing and subtyping HMs morphologically. On the other hand, one could also imagine a diminished use of the moderator code due to increased use of various ancillary techniques like immunohistochemical analyses and molecular genotyping.[26,27] In our study, the moderator was added to the first-time HM code for 15% of the women in the DPR, and there was no significant difference regarding the morphological subdiagnosis of mole or registration before and after 2004. Others also use diagnoses of conceptuses suspected of HMs that cannot be classified unequivocally.

In one study conceptuses suspected of PHM, but with villus morphology insufficient for a definitive diagnosis of PHM, were classified as either "morphology abnormal, favor PHM" or "morphology abnormal, probable nonmolar, PHM not excluded".[28] Thus, there is a continuing need for categories for classifying "inconclusive" and "suspected cases". The observation that disregarding women registered with the moderator code ÆYYY00 in the DPR caused an increased completeness and a reduced PPV of registration in the NPR and DCR indicates that some of the conceptuses classified by the pathologist as a suspected HM were regarded by the clinicians to be HMs whereas others were regarded to be nonmolar conceptuses. Understanding the meaning of the moderator code in the SNOMED coding system, the researcher can decide if it is most informative to include or exclude women registered with this code in the DPR in a given study.

In this study, data retrieval from the DPR revealed some challenges in the Danish SNOMED regarding the morphology code M91000 in the period examined. This code indicates that the women were diagnosed with an HM. Since the code M91030 indicates that the women were diagnosed with a PHM, one could assume the code M91000 was intended for women with CHMs. However, an HM, not otherwise specified, could also correctly have been encoded M91000. Thus, we do not know which of the women with this code had a CHM, and which had an HM that was not (or could not be) specified.

A good coding system should allow classification in exhaustive and mutually exclusive categories. We only disclosed that the codes for the morphological subdiagnoses in the DPR were not mutually exclusive during this study almost 25 years after the codes had been introduced, illustrating that the quality of registries can benefit from the use of the data. In January 2014, the codes for HM in the DPR were adjusted by introducing specific morphology codes for "CHM" and "HM, not otherwise specified".[21] Thus, it would be relevant to repeat the validation study in 5–10 years.

Conclusion

For studying frequencies of HMs in the period 1999–2009, the DPR appears to be the most valid source. For a given study, one may wish to include or exclude women diagnosed with a suspected HM. The NPR and the DCR appear to be equally valid relative to the DPR before 2004. From 2004 onward, the completeness and PPV for the DCR were higher than for the NPR.

Acknowledgment

Disclosure

The authors report no conflicts of interest in this work.

References

1. Wells M. The pathology of gestational trophoblastic disease: recent advances. *Pathology*. 2007;39(1):88–96.
2. Altieri A, Franceschi S, Ferlay J, Smith J, La Vecchia C. Epidemiology and aetiology of gestational trophoblastic diseases. *Lancet Oncol*. 2003;4(11):670–678.
3. Brown J, Naumann RW, Seckl MJ, Schink J. 15 years of progress in gestational trophoblastic disease: scoring, standardization, and salvage. *Gynecol Oncol*. 2017;144(1):200–207.
4. Joneborg U, Folkvaljon Y, Papadogiannakis N, Lambe M, Marions L. Temporal trends in incidence and outcome of hydatidiform mole: a retrospective cohort study. *Acta Oncol*. 2018;0(0):1–6.
5. Eysbouts YK, Bulten J, Ottevanger PB, et al. Trends in incidence for gestational trophoblastic disease over the last 20 years in a population-based study. *Gynecol Oncol*. 2016;140(1):70–75.
6. Matsui H, Kihara M, Yamazawa K, Mitsuhashi A, Seki K, Sekiya S. Recent changes of the incidence of complete and partial mole in Chiba prefecture. *Gynecol Obstet Invest*. 2007;63(1):7–10.
7. Salehi S, Eloranta S, Johansson AL, Bergström M, Lambe M. Reporting and incidence trends of hydatidiform mole in Sweden 1973-2004. *Acta Oncol*. 2011;50(3):367–372.
8. Flam F, Rutqvist LE. Under-registration of gestational trophoblastic disease in the Swedish Cancer Registry. *Eur J Epidemiol*. 1992;8(5):683–686.
9. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(7 Suppl):30–33.
10. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health*. 2011;39(7 Suppl):42–45.
11. Erichsen R, Lash TL, Hamilton-Dutoit SJ, Bjerregaard B, Vyberg M, Pedersen L. Existing data sources for clinical epidemiology: the Danish National Pathology Registry and Data Bank. *Clin Epidemiol*. 2010;2:51–56.
12. Andersen TF, Madsen M, Jørgensen J, Mellemkjoer L, Olsen JH. The Danish National Hospital Register. A valuable source of data for modern health sciences. *Dan Med Bull*. 1999;46(3):263–268.
13. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449–490.
14. Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry--history, content, quality and use. *Dan Med Bull*. 1997;44(5):535–539.
15. Bjerregaard B, Larsen OB. The Danish Pathology Register. *Scand J Public Health*. 2011;39(7 Suppl):72–74.
16. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health*. 2011;39(7 Suppl):22–25.
17. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol*. 2014;29(8):541–549.
18. Sebire NJ. Histopathological diagnosis of hydatidiform mole: contemporary features and clinical implications. *Fetal Pediatr Pathol*. 2010;29(1):1–16.
19. Savage JL, Maturen KE, Mowers EL, et al. Sonographic diagnosis of partial versus complete molar pregnancy: a reappraisal. *J Clin Ultrasound*. 2017;45(2):72–78.
20. Fowler DJ, Lindsay I, Seckl MJ, Sebire NJ. Routine pre-evacuation ultrasound diagnosis of hydatidiform mole: experience of more than 1000 cases from a regional referral center. *Ultrasound Obstet Gynecol*. 2006;27(1):56–60.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
22. Vang R, et al. Diagnostic reproducibility of hydatidiform moles: ancillary techniques (p57 immunohistochemistry and molecular genotyping) improve morphologic diagnosis. *Am J Surg Pathol*. 2012;36(3):443–453.
23. Buza N, Hui P. Partial hydatidiform mole: histologic parameters in correlation with DNA genotyping. *Int J Gynecol Pathol*. 2013;32(3):307–315.
24. Sebire NJ, Makrydimas G, Agnantis NJ, Zagorianakou N, Rees H, Fisher RA. Updated diagnostic criteria for partial and complete hydatidiform moles in early pregnancy. *Anticancer Res*. 2003;23(2C):1723–1728.
25. Sebire NJ, Lindsay I. Current issues in the histopathology of gestational trophoblastic tumors. *Fetal Pediatr Pathol*. 2010;29(1):30–44.
26. Ronnett BM, Descipio C, Murphy KM. Hydatidiform moles: ancillary techniques to refine diagnosis. *Int J Gynecol Pathol*. 2011;30(2):101–116.
27. Banet N, Descipio C, Murphy KM, et al. Characteristics of hydatidiform moles: analysis of a prospective series with p57 immunohistochemistry and molecular genotyping. *Mod Pathol*. 2014;27(2):238–254.
28. Fisher RA, Tommasi A, Short D, Kaur B, Seckl MJ, Sebire NJ. Clinical utility of selective molecular genotyping for diagnosis of partial hydatidiform mole; a retrospective study from a regional trophoblastic disease unit. *J Clin Pathol*. 2014;67(11):980–984.