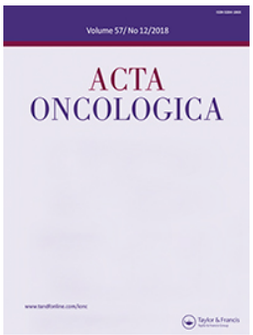

Taylor & Francis
Taylor & Francis Group

Acta Oncologica

ISSN: 0284-186X (Print) 1651-226X (Online) Journal homepage: www.informahealthcare.com/journals/ionc20

Low proportion of unreported cervical treatments in the cancer registry of Norway between 1998 and 2013

Maarit K. Leinonen, Svenn A. Hansen, Gry Baadstrand Skare, Inger Berit Skaaret, Monica Silva, Tom Børge Johannesen & Mari Nygård

To cite this article: Maarit K. Leinonen, Svenn A. Hansen, Gry Baadstrand Skare, Inger Berit Skaaret, Monica Silva, Tom Børge Johannesen & Mari Nygård (2018) Low proportion of unreported cervical treatments in the cancer registry of Norway between 1998 and 2013, Acta Oncologica, 57:12, 1663-1670, DOI: 10.1080/0284186X.2018.1497296

To link to this article: https://doi.org/10.1080/0284186X.2018.1497296

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 31 Aug 2018.

Submit your article to this journal

Article views: 1138

View related articles

View Crossmark data

Citing articles: 4 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.informahealthcare.com/action/journalInformation?journalCode=ionc20

ACTA ONCOLOGICA
2018, VOL. 57, NO. 12, 1663–1670
https://doi.org/10.1080/0284186X.2018.1497296

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Low proportion of unreported cervical treatments in the cancer registry of Norway between 1998 and 2013

Maarit K. Leinonen[a], Svenn A. Hansen[b], Gry Baadstrand Skare[c], Inger Berit Skaaret[c], Monica Silva[d], Tom Børge Johannesen[d] and Mari Nygård[a]

[a]Department of Research, Cancer Registry of Norway, Oslo, Norway; [b]Department of Health Management and Health Economics, University of Oslo, Oslo, Norway; [c]Unit of Cervical Cancer Screening, Cancer Registry of Norway, Oslo, Norway; [d]Department of Registration, Cancer Registry of Norway, Oslo, Norway

ABSTRACT

Background: Accurate information about treatment is needed to evaluate cervical cancer prevention efforts. We studied completeness and validity of reporting cervical treatments in the Cancer Registry of Norway (CRN).

Material and methods: We identified 47,423 (92%) high-grade cervical dysplasia patients with and 3983 (8%) without recorded treatment in the CRN in 1998–2013. We linked the latter group to the nationwide registry of hospital discharges in 1998–2015. Of patients still without treatment records, we randomly selected 375 for review of their medical history. Factors predicting incomplete treatment records were assessed by multiple imputation and logistic regression.

Results: Registry linkage revealed that 10% (401/3983) of patients received treatment, usually conization, within one year of their initial high-grade dysplasia diagnosis. Of those, 11% ($n = 44$) were missing due to unreporting and 89% ($n = 357$) due to misclassification at the CRN. Of all cases in medical review, patients under active surveillance contributed almost 60% (223/375). Other reasons of being without recorded treatment were uncertain dysplasia diagnosis, invasive cancer or death. Coding error occurred in 19% (73/375) of randomly selected cases. CRN undercounted receipt of treatment by 38% ($n = 1526$) among patients without recorded treatment which translates into 97% overall completeness of treatment data. Incomplete treatment records were particularly associated with public laboratories, patients aged 40–54 years, and the latest study years.

Conclusions: CRN holds accurate information on cervical treatments. Completeness and particularly validity can be further improved through the establishment of new internal routines and regular linkage to hospital discharges.

ARTICLE HISTORY
Received 23 February 2018
Accepted 29 June 2018

Introduction

Cervical cancer control relies on early detection and treatment of preinvasive lesions [1]. Whenever abnormal cells are detected in a cervical smear, a thorough evaluation with colposcopy-directed punch biopsies follows. When appropriate, treatment eradicates a preinvasive lesion preventing its further progression into invasive cancer, and a success rate is over 90% [2]. Two types of treatment are available: excisional and ablative (destructive) procedures. Excisional techniques include most commonly used loop electrosurgical excision (LEEP), cold-knife conization, laser conization and in some cases hysterectomy. Ablative or destructive procedures include laser ablation, cryotherapy, radical diathermy and cold coagulation. Local ablation or destruction can be considered if pretreatment biopsies show no evidence of invasive or glandular disease, and the entire transformation zone is visible [3]. In Norway, almost 99% of cervical preinvasive lesions are treated with excisional procedures, usually with LEEP (46%) or laser conization (38%). Ablative procedures constitute about 1% of treatments only [4].

Registration of precancerous lesions of cervix uteri is common in the Nordic cancer registries [5]. While the completeness of vulvar and vaginal preinvasive lesions in Norway has been formally evaluated [6], less is known about registration of cervical lesions. It has been estimated that the CIN Registry has more than 80% of ascertainment of treatments of cervical precancerous lesions [4,7]. Accurate information about treatment of cervical lesions is needed to evaluate cervical cancer prevention efforts and burden on the health care system. To determine the extent of missing registry data, administrative claims linked to registry data provide a robust tool for quality evaluation [8]. We studied completeness and validity of reporting treatments of

CONTACT Maarit K. Leinonen maarit.leinonen@kreftregisteret.no Department of Research, Cancer Registry of Norway, Ullernchausseen 64, Oslo 0379, Norway

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

cervical lesions at the Cancer Registry of Norway (CRN) for the first time. We also made a descriptive analysis on reasons of being without recorded treatment for high-grade cervical dysplasia.

Material and methods

The CRN holds information on cervical treatments in three databases: the Incidence Database, the Histology Registry and the CIN Registry. Reporting is mandatory and is based on pathology and clinical notifications (messages). The Incidence Database (established 1951) provides information about site, histology and stage of all new cancer cases while the Histology Registry (established 2002) provides information on all histology specimens, including normal findings, taken from a cervix. The main purpose of the Incidence Database and the Histology Registry is to capture all cervical diagnoses while the CIN Registry (established 1997) holds information on preoperative findings, histology and cervical treatment procedures. Reporting of histology during the study period was based on modified version of the SNOMED coding system [9]. Histology was classified according to the SNOMED system in the Histology Registry and ICD-O-3 [10] in the Incidence Database.

Identification of patients diagnosed with a high-grade dysplasia, i.e., cervical intraepithelial lesion grade 2 (CIN 2),

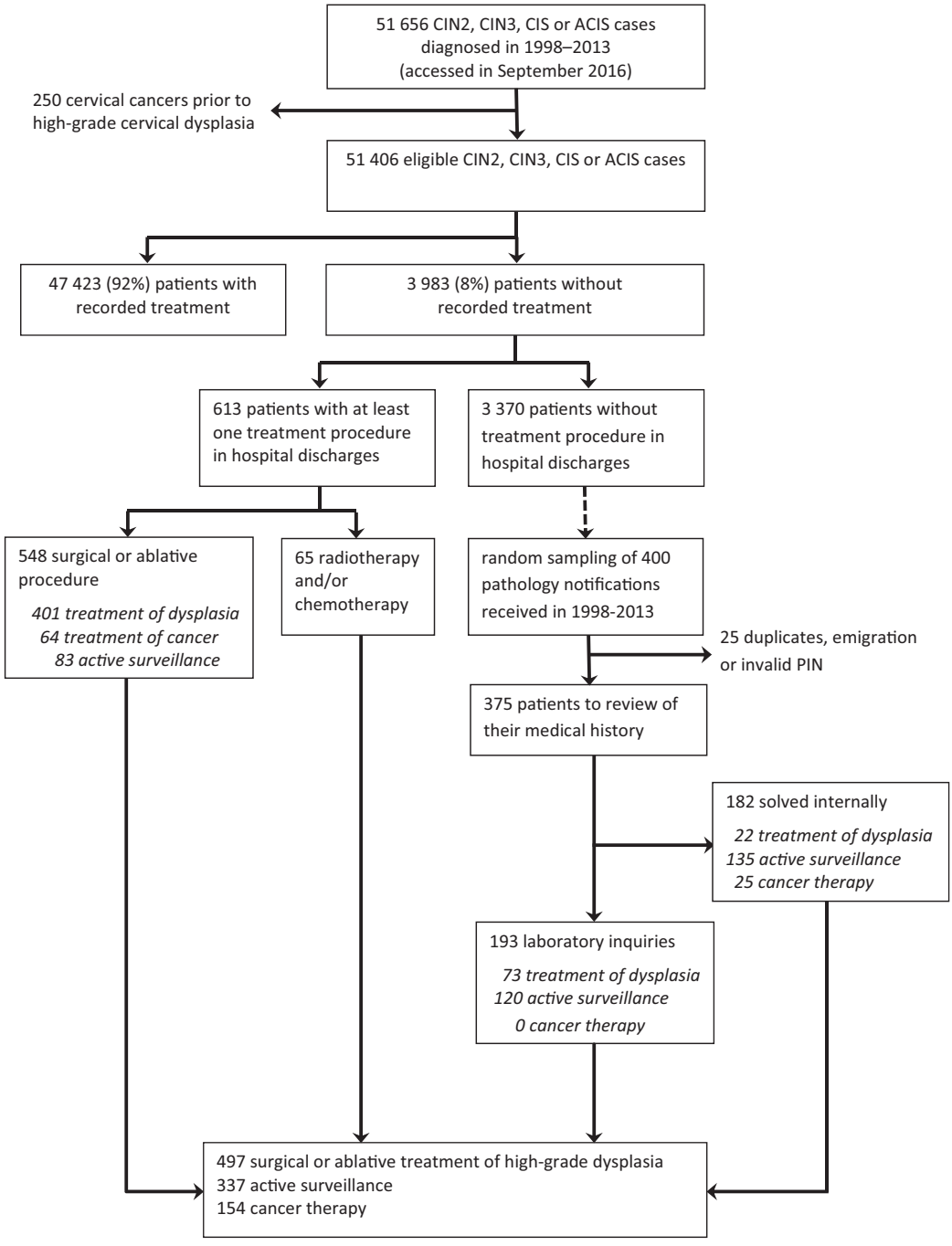

Figure 1. Identification of 51,406 high-grade cervical dysplasia patients diagnosed in Norway between 1998 and 2013 and their assignment to treatment groups in the beginning and during the study.

cervical intraepithelial lesion grade 3 (CIN 3), carcinoma in situ (CIS) or adenocarcinoma in situ (ACIS), and their assignment to treatment groups is given in Figure 1. We extracted all high-grade dysplasias diagnosed between 1 January 1998 and 31 December 2013 and registered in any of the three database in September 2016 ($n = 51{,}656$). We excluded 250 patients who had been diagnosed with invasive cervical cancer before the high-grade cervical dysplasia. We categorized patients into two mutually exclusive groups: 'with recorded treatment' and 'without recorded treatment'. A patient was considered of being without treatment if there was no recorded information on excisional or ablative treatment after their initial high-grade dysplasia diagnosis.

We then linked 3983 (7.7%) patients with high-grade dysplasia and without recorded treatment to hospital discharges recorded in the Patient Administrative Data System (PAS, available from 1998 to 2010) and the Norwegian Patient Registry (NPR, available from 2010 to 2015). We used the unique personal identification number (PIN) that all Norwegian residents have, and requested all hospitalizations and outpatient visits in these nationwide registries, recorded with the ICD-10 diagnostic codes of interest (C53, D06-D07, N87, N89), and relevant medical, surgical and radiological procedures that were recorded irrespective of ICD-10 code. We then categorized found treatment procedures into three categories: surgical or ablative treatment of high-grade dysplasia, active surveillance (see below) and cancer therapy.

Among 3370 patients with high-grade dysplasia but no treatment procedure in hospital discharges, we took a random sample in order to review their medical history in detail. We made random sampling of 400 pathology notifications received in the CRN in 1998–2013. We excluded 25 patients because of duplicate sampling, emigration after high-grade dysplasia or invalid PIN leaving 375 eligible patients for medical review.

First, we reviewed records of these 375 patients from all three databases to identify any in-house discrepancies in coding treatment data. We solved internally 49% of cases ($n = 182$), and these were not subjected to laboratory inquiries, i.e., to external review. Remaining 193 patients were ordered according to the name of the laboratory who reported the high-grade dysplasia diagnosis. Laboratory inquiries were done by regular mail or occasionally by phone when there were only a few uncertain cases per laboratory. We provided the laboratories with diagnosis, date of diagnosis and time period of interest for which we asked them to send us all relevant patient records for all listed patients. Laboratories received up to one reminder letter if they did not respond within 3–4 weeks after the first inquiry.

We compared information on treatment in the external data sources to the registry data in the CRN. Treatment provided within one year after the initial diagnosis was considered as a primary (first-line) treatment of which the CRN should hold records. Treatment given more than one year after high-grade dysplasia diagnosis was categorized as active surveillance also called as watchful waiting as a primary treatment. Active surveillance also included women to whom (1) the CRN had received a clinical notification explicitly saying no treatment, (2) the CRN had received a pathology notification from the same hospital visit that confirmed punch biopsy and (3) the CRN hold information on colposcopy, ultrasound, Pap smear or similar diagnostic procedures following the high-grade dysplasia indicating that clinical management strategy was conservative.

Statistical analyses

We made a descriptive analysis on reasons for unregistered treatment and performed a non-parametric test of trend (P_{trend}) to study possible trends across groups in these reasons where appropriate. We also studied factors that predicted incomplete treatment records in the CRN using logistic regression. Results from the data linkage and from medical review of randomly selected patients were imputed to the whole population without recorded treatment ($n = 3983$) using a multiple imputation by chained equations (MICE) approach. Incomplete treatment records in the CRN were our outcome variable and it was assumed to be missing at random [11]. Explanatory variables included age group at diagnosis in three categories (less than 40, 40–54 and 55 years and above), severity of diagnosis (CIN 2 vs. CIN 3/CIS/ACIS), type of laboratory (university hospital, other hospital, private or unknown) and time period (years 1998–2003, 2004–2008 and 2009–2013). In addition, interaction between severity of diagnosis and age group, exact name of laboratory and exact time since high-grade dysplasia in years were included in the imputation models as possible predictors of missing information. We created 40 imputed data sets of which estimates were combined using Rubin's rules [12,13]. Statistical analyses were done using Stata version 15.0 (StataCorp, College Station, TX, USA). Ethical approval was not needed as the CRN has an obligation to continuously monitor and measure the data quality of its own databases.

Results

Registry linkage revealed that 14% (548 out of 3983) of patients without recorded treatment had a surgical or ablative procedure performed, and 2% ($n = 65$) received radiotherapy- and/or chemotherapy within one year after high-grade dysplasia diagnosis (Figure 1). Fifteen percent of identified treatment procedures (83 out of 548) were considered as under active surveillance, typically, because treatment was given more than one year after high-grade dysplasia diagnosis (Table 1). Only seven patients had an erroneous procedural code in hospital discharges and a pathology report in the CRN from the same hospital visit confirmed biopsy. Forty-nine per cent of primary treatments (303 out of 613) were conizations, 22% were hysterectomies and 5% ablative procedures. Hysterectomy was associated with increasing age ($P_{trend} = .000$) and a glandular lesion (ACIS, $P_{trend} = .000$), whereas conization was the most prevalent treatment type in women of reproductive age and in squamous cell lesions. The great majority of identified treatments were among patients who had been diagnosed with high-grade dysplasia during the last time period in 2009–2013. The CRN had

Table 1. Results from linkage of 3983 patients without recorded treatment for high-grade cervical dysplasia in the CRN to the nationwide registry of hospital discharges in 1998–2015.

		Primary treatment type within ≤365 days				
	Patients *n* (% of total)	Active surveillance[a] *n* (% of total)	Radio- and/or chemotherapy *n* (% of total)	Conization *n* (% of total)	Hysterectomy *n* (% of total)	Ablative procedure *n* (% of total)
Age group at diagnosis						
<40 years	286 (46.7%)	44 (53.0%)	12 (18.5%)	193 (63.7%)	12 (9.0%)	25 (89.3%)
40–54 years	164 (26.8%)	24 (28.9%)	20 (30.8%)	78 (25.7%)	42 (31.1%)	0 (–)
≥55 years	163 (26.6%)	15 (18.1%)	33 (50.8%)	32 (10.6%)	80 (59.7%)	3 (10.7%)
P_{trend}		.080	.000	.000	.000	.000
Severity of diagnosis						
CIN 2	197 (32.1%)	47 (56.6%)	1 (1.5%)	99 (32.7%)	32 (23.9%)	18 (64.3%)
CIN 3, CIS	348 (56.8%)	32 (38.6%)	58 (89.2%)	198 (65.4%)	50 (37.3%)	10 (35.7%)
ACIS	68 (11.1%)	4 (4.8%)	6 (9.2%)	6 (2.0%)	52 (38.8%)	0 (–)
P_{trend}		.000	.000	.000	.000	.000
Laboratory of diagnosis						
University hospital	285 (46.5%)	34 (41.0%)	20 (30.8%)	141 (46.5%)	77 (57.5%)	13 (46.4%)
Other hospital	287 (46.8%)	39 (47.0%)	34 (52.3%)	149 (49.2%)	51 (38.1%)	14 (50.0%)
Private or unknown	41 (6.7%)	10 (12.0%)	11 (16.9%)	13 (4.3%)	6 (4.5%)	1 (3.6%)
Time of diagnosis						
1998–2003	47 (7.7%)	22 (26.5%)	5 (7.7%)	1 (0.3%)	19 (14.2%)	0 (–)
2004–2008	60 (9.8%)	16 (19.3%)	10 (15.4%)	1 (0.3%)	33 (24.6%)	0 (–)
2009–2013	506 (82.5%)	45 (54.2%)	50 (76.9%)	301 (99.3%)	82 (61.2%)	28 (100.0%)
P_{trend}		.000	.410	.000	.000	.000
Total	613 (100.0%)	83 (100.0%)	65 (100.0%)	303 (100.0%)	134 (100.0%)	28 (100.0%)

[a]Pathology notification from the same hospital visit confirmed biopsy or treatment was given more than one year after high-grade dysplasia diagnosis.

Table 2. Patients without recorded treatment for high-grade cervical dysplasia in the CRN who were solved internally by age, severity, laboratory type and time period of diagnosis ($n = 182$).

		Reasons of being without recorded treatment for high-grade cervical dysplasia				
	Patients *n* (% of total)	Cervical cancer ≤6 months (prevalent) *n* (% of total)	Other cancer or death ≤12 months *n* (% of total)	Active surveillance[a] *n* (% of total)	Uncertain dysplasia diagnosis *n* (% of total)	Uncoded treatment *n* (% of total)
Age group at diagnosis						
<40 years	118 (64.8%)	1 (11.1%)	3 (18.8%)	79 (76.7%)	22 (68.8%)	13 (59.1%)
40–54 years	23 (12.6%)	2 (22.2%)	2 (12.5%)	9 (8.7%)	4 (12.5%)	6 (27.3%)
≥55 years	41 (22.5%)	6 (66.7%)	11 (68.8%)	15 (14.6%)	6 (18.8%)	3 (13.6%)
P_{trend}		.000	.000	.000	.566	.851
Severity of diagnosis						
CIN 2	107 (58.8%)	0 (–)	4 (25.0%)	71 (68.9%)	24 (75.0%)	8 (36.4%)
CIN 3, CIS	64 (35.2%)	8 (88.9%)	10 (62.5%)	30 (29.1%)	6 (18.8%)	10 (45.5%)
ACIS	11 (6.0%)	1 (11.1%)	2 (12.5%)	2 (1.9%)	2 (6.3%)	4 (18.2%)
P_{trend}		.001	.006	.000	.102	.005
Laboratory of diagnosis						
University hospital	84 (46.2%)	4 (44.4%)	5 (31.3%)	46 (44.7%)	14 (43.8%)	15 (68.2%)
Other hospital	83 (45.6%)	5 (55.6%)	7 (43.8%)	50 (48.5%)	14 (43.8%)	7 (31.8%)
Private or unknown	15 (8.2%)	0 (–)	4 (25.0%)	7 (6.8%)	4 (12.5%)	0 (–)
Time of diagnosis						
1998–2003	36 (19.8%)	3 (33.3%)	7 (43.8%)	18 (17.5%)	4 (12.5%)	4 (18.2%)
2004–2008	51 (28.0%)	2 (22.2%)	6 (37.5%)	24 (23.3%)	12 (37.5%)	7 (31.8%)
2009–2013	95 (52.2%)	4 (44.4%)	3 (18.8%)	61 (59.2%)	16 (50.0%)	11 (50.0%)
P_{trend}		.404	.002	.068	.687	.970
Total	182 (100.0%)	9 (100%)	16 (100%)	103 (100%)	32 (100.0%)	22 (100%)

[a]Colposcopy, ultrasound, Pap smear or other diagnostic procedures following the high-grade dysplasia or a clinical notification explicitly saying no treatment.

complete data on radio- and/or chemotherapies administered for cervical cancer.

Among internally reviewed 182 cases, the most common reason of unregistered treatment was active surveillance (Table 2). These patients contributed 57% of all cases. Active surveillance was particularly observed among patients younger than 40 years (P_{trend}=.000) and among patients with less severe lesion (CIN 2, P_{trend}=.000). Eighteen per cent of internally reviewed cases had an uncertain dysplasia diagnosis in the first place. Their pathology reports stated CIN 1-2 or ungraded CIN where moderate dysplasia cannot be ruled out. These were considered to be managed as CIN 1 lesions which do not require treatment. Nine patients (5%) developed an invasive cervical cancer within six months. Of these, eight patients were treated with radiotherapy and one had refused treatment. Nine per cent of patients had some other cancer or died within one year after the high-grade dysplasia diagnosis.

Among cases subjected to external review, we received at least one cytology or histology report following the high-grade dysplasia diagnosis for 179 patients. For 14 patients (7% of all inquired) no any records were identified from the laboratories we contacted for the asked time period. These patients were assumed to have complete records at the CRN

Table 3. Patients without recorded treatment for high-grade cervical dysplasia in the CRN who were inquired from laboratories by age, severity, laboratory type and time period of diagnosis ($n = 193$).

		Reasons of being without recorded treatment for high-grade cervical dysplasia		
	Patients *n* (% of total)	Active surveillance *n* (% of total)	Unreported treatment *n* (% of total)	Uncoded treatment *n* (% of total)
Age group at diagnosis				
<40 years	138 (71.5%)	92 (76.7%)	14 (63.6%)	32 (62.8%)
40–54 years	34 (17.6%)	17 (14.2%)	5 (22.7%)	12 (23.5%)
≥55 years	21 (10.9%)	11 (9.2%)	3 (13.6%)	7 (13.7%)
P_{trend}		.070	.435	.154
Severity of diagnosis				
CIN 2	96 (49.7%)	65 (54.2%)	6 (27.3%)	25 (49.0%)
CIN 3, CIS	95 (49.2%)	54 (45.0%)	16 (72.7%)	25 (49.0%)
ACIS	2 (1.0%)	1 (0.8%)	0 (–)	1 (2.0%)
P_{trend}		.114	.041	.793
Laboratory of diagnosis				
University hospital	85 (44.0%)	52 (43.3%)	8 (36.4%)	25 (49.0%)
Other hospital	68 (35.2%)	44 (36.7%)	8 (36.4%)	16 (31.4%)
Private or unknown	40 (20.7%)	24 (20.0%)	6 (27.3%)	10 (19.6%)
Time of diagnosis				
1998–2003	71 (36.8%)	40 (33.3%)	16 (72.7%)	15 (29.4%)
2004–2008	77 (39.9%)	46 (38.3%)	2 (9.1%)	29 (56.9%)
2009–2013	45 (23.3%)	34 (28.3%)	4 (18.2%)	7 (13.7%)
P_{trend}		.049	.007	.810
Total	193 (100.0%)	120 (100%)	22 (100%)	51 (100%)

and considered under active surveillance group without any screening or diagnostic follow-up. All results from laboratory inquiries are shown in Table 3. Overall, 62% of patients (120 out of 193) untreated for high-grade dysplasia were under active surveillance, and their proportion was highest during the latest study years (P_{trend}=.049).

Coding error in the CRN had occurred in 19% of randomly selected cases (Tables 2 and 3). Proportion of coding errors was 12% in internal review and 26% in external review. Of those internally reviewed cases, one was destructive treatment of vaginal CIS, two were hysterectomies with normal findings in cervix and 19 were conizations reported using an unspecific SNOMED code (T83000 Cervix Uteri UNS). Two-thirds of coding errors in external review were reported using an unspecific SNOMED code and one-third was misclassification of treatment status in the CRN albeit of accurate reporting. Twenty-two cases (6% of all medically reviewed) were due to previously unreported treatment to the CRN.

Finally, we categorized results from the PAD and NPR linkage to comply with reasons of unregistered treatment from medical review. Almost half of identified hysterectomies ($n = 63$) were treatments against other gynaecological cancer than cervical cancer, usually endometrial cancer. Two patients had ovarian cancer and one patient had both endometrial and ovarian cancer. One conization was found to be an erroneous code in the NPR and the CRN confirmed treatment as a radical hysterectomy against endometrial cancer. Thus, 401 surgical or ablative procedures were previously unregistered treatments against high-grade dysplasia. Of these, 11% ($n = 44$) were due to unreporting and 89% ($n = 357$) due to misclassification. Ablative procedures constituted 58% of all unreported treatments. Combined dataset for imputation included 497 surgical or ablative treatments of high-grade dysplasia, 337 patients with active surveillance and 154 patients with cancer therapy.

Imputation models revealed that the CRN undercounted receipt of surgical or ablative procedure by 38% in the whole population without recorded treatment (1526 out of 3983) (Table 4). This would indicate that completeness of surgical or ablative procedures for cervical high-grade dysplasia in the CRN overall is 97% (47,423/48,949). Results from logistic regression showed that incomplete treatment records in the CRN were particularly associated with public laboratories, with the most recent study years and with patients aged from 40 to 54 years. Severity of disease also predicted incomplete treatment records so that treatment for CIN 3, CIS or ACIS was more often incomplete than treatment for CIN 2.

Discussion

This was the first time completeness and validity of reporting treatment for cervical high-grade dysplasia in the CRN was thoroughly evaluated. We found that CRN undercounted receipt of surgical or ablative procedure by 38% among women without recorded treatment. This translates into 97% overall completeness of treatment data. Reporting using unspecific SNOMED codes was more prominent threat to data quality than underreporting. Completeness and particularly validity can be further improved through establishing new internal routines and regular linkage between the NPR and the CRN.

Patients under active surveillance contributed 15% of cases in a register linkage and almost 60% cases in a random sample. Active surveillance together with ablative procedures were particularly attributed to patients younger than 40 years and to CIN 2 lesions. Awareness that particularly deep excisions and repeat treatments increase the risk of prematurity and perinatal complications has increased [4,14]. Ablative therapies have minor impact on adverse obstetric outcomes and hence may be alternatives in reproductive-

Table 4. Predictors of incomplete treatment records among 3983 high-grade dysplasia patients without recorded treatment in the CRN.

	Incomplete treatment records			
	Observed data $n = 497$ n (% of total)	Imputed data $n = 1029$ n (% of total)	Completed $n = 1526$ n (% of total)	Multivariate odds ratio (95% CI)
Age group at diagnosis				
<40 years	286 (57.6%)	673 (65.4%)	959 (62.8%)	Ref.
40–54 years	137 (27.6%)	232 (22.5%)	369 (24.2%)	3.23 (1.99–5.24)
≥55 years	74 (14.9%)	124 (12.1%)	198 (13.0%)	1.20 (0.61–2.34)
Severity of diagnosis				
CIN 2	178 (35.8%)	479 (46.6%)	657 (43.1%)	Ref.
CIN 3, CIS, ACIS	319 (59.6%)	550 (53.4%)	869 (56.9%)	3.41 (2.53–4.59)
Laboratory of diagnosis				
University hospital	242 (48.7%)	562 (54.6%)	804 (52.7%)	Ref.
Other hospital	222 (44.7%)	349 (33.9%)	571 (37.4%)	0.83 (0.63–1.08)
Private or unknown	33 (6.6%)	118 (11.5%)	151 (9.9%)	0.40 (0.23–0.68)
Time period of diagnosis				
1998–2003	43 (8.7%)	318 (30.9%)	361 (23.7%)	0.29 (0.19–0.44)
2004–2008	49 (9.9%)	278 (27.0%)	327 (21.4%)	0.28 (0.21–0.38)
2009–2013	405 (81.5%)	433 (42.1%)	838 (54.9%)	Ref.

aged women [15]. Given the high proportion of CIN 2 lesions that go into spontaneous regression, watchful waiting has become an attractive option for young women [15–21]. Our study showed that albeit national guidelines, where CIN 2 typically is treated, active surveillance plays a role in clinical practice. However, results have to be interpreted keeping in mind that CIN 2 is an equivocal diagnosis of poor reproducibility [22,23] making it a challenging diagnosis for registration and time trend analysis. Also, current coding routines in the CRN should be adapted to capture active surveillance as an option for treatment.

Incomplete treatment records in the CRN were particularly associated with the latest study years. There is always a trade-off between timely data and the extent to which data are complete and accurate. Complete information on the first course of treatment may not be available until one year after the initial diagnosis, and also reporting delays exist. While there are no internal guidelines for timelines, within two years of the close of the diagnosis year, at least 95% completeness is usually expected [24]. Our last year of diagnosis was 2013. Still, 88 cases without recorded treatment at the time of data extraction had correct information 6–12 months later at the time of the review (data not shown). This means that the latest years of diagnosis may not be the most optimal setting to study treatment pathways.

While the CIN Registry is routinely cross-linked to the Histology Registry to check for missing information on treatment, there is no routine to verify treatment data in these against the Incidence Database. Our study revealed that the Incidence Database is an important data source to complement the treatment data and particularly to decide about the primary site of the tumor. Using the Histology Registry and CIN Registry alone for statistics and research will overestimate the proportion of untreated patients and provide higher incidence of ACIS in the population. This is because of reporting using unspecific SNOMED codes and because screening program does not receive systematic information on hysterectomies. Whether simultaneous cervical ACIS and endometrial cancer constitute one or two incident cases may have consequences when evaluating incidence patterns over time and between countries. Endometrial cancer is the most common gynecologic cancer among Norwegian women [25], and prevalence of its major risk factors, namely obesity and overweight, is increasing rapidly [26,27]. Lack of uniform classification and analysis of multiple primaries is a known problem related to a cancer registry for all tumors versus any single domain oriented registry [28]. Recently, health research is increasingly adapting new analytical techniques to identify patterns in large data sets. Differences in the way data are collected and stored (metadata) and inconsistencies between data sources can reduce value of applying machine learning and data mining [29–31]. Therefore, institutions collecting health data should take these new opportunities into account when organizing their data and supporting metadata [29].

Hospital discharges are a quick and feasible way to study resource utilization. In research using hospital discharges, use of diagnostic and procedure codes is common but infrequently validated. Patients registered on a code frequently do not have the condition it represents [32,33]. We scrutinized thoroughly identified procedural codes and found only a handful of erroneous procedural codes in the PAS and the NPR. Also agreement between diagnostic codes in the NPR and the CRN has previously shown to be high for the six most common cancer types [34]. Therefore, we do not anticipate our results to be largely affected by erroneous codes. In fact, the NPR is already a key source in finding information on unreported cancer cases [25]. Proportion of unreported high-grade dysplasia cases overall was low in our study. This confirms that data in the CRN can be used to monitor incidence of all Human Papillomavirus-related preinvasive gynecological lesions with high precision [6]. Furthermore, our study demonstrates the value of supplementing registry data on treatment by linkage to hospital discharges or administrative claims when these are of high quality, and similar results have been reported previously [8].

Cancer registries often lack data on cancer therapies, especially those administered in the outpatient setting such as chemotherapy, endocrine therapy and radiation [8,35-37]. Even the information on first-line treatment is far from

perfect [5]. Not surprisingly, European cancer registries' contribution to cancer care outcomes is still rather limited [38]. We found that all cervical cancer cases had information on administered radiotherapy and occasionally also on chemotherapy. This substantial result is most likely attributed to access to data from radiation machines since 1997 [5]. Currently, the information on chemotherapy in the CRN is on a dichotomous level except for nine quality registries. Data on chemotherapy is not easy to supplement from administrative claims as medical procedures were often recorded without any ICD-10 code. This makes it hard to distinguish between primary treatment and treatment for recurrence/metastasis/other cancer in case of multiple malignancies.

We included several covariates associated with patient, time, service provider and their interactions that could relate to missingness into imputation models. Therefore, we feel that missing at random assumption was reasonable. Furthermore, MICE was a correct approach since hospital discharges were not equally available for the whole study period. This was clearly seen in the data linkage where the majority of identified procedures were among patients diagnosed with high-grade dysplasia in 2009–2013. Complete case analysis emphasized data from registry linkage and lead to somewhat different estimates. Furthermore, it is possible that private laboratories do not report to the CRN and to the hospital discharges with the same intensity as public laboratories. A study comparing cancer diagnoses in the NPR and in the CRN found that the private sector rarely reported patient's PIN to the NPR which limits possibilities to further use this data [34]. A study from New Zealand clearly pointed toward a non-reporting private hospital 'effect' on completeness of cancer surgeries [37]. We noticed slight differences in responses to our inquiries between public and private laboratories. Private laboratories tended to provide fewer cytology and pathology notifications per patient which could be due to limited number of visits to clinician, but may also indicate less intensive collection of patient records.

A limitation of our study is that we only studied patients without recorded treatment, and we do not know about validity of patients with recorded treatment. Therefore, our estimates about overall completeness of 97% should be interpreted with some caution. Still, our study indicates that the CRN holds accurate information about treatment for cervical high-grade dysplasia and also for cervical cancer. This makes the CRN valuable data source to evaluate cervical cancer prevention and treatment efforts in the population.

Conclusions

CRN holds accurate information on cervical treatments. CRN is a valuable data source to evaluate cervical cancer prevention and treatment efforts in the population. The best data for statistics and research is obtained by combining information from all relevant databases. Completeness and particularly validity of each database individually can be further improved through the establishment of new internal routines and regular linkage to hospital discharges.

Acknowledgments

We thank Suzanne Campbell, Espen Enerly and Lena Holmström at the Cancer Registry of Norway and all doctors and Medical Laboratory Technologists at receiving laboratories for their collaboration and contribution.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Maarit K. Leinonen http://orcid.org/0000-0002-7631-4749

References

[1] Von Karsa L, Arbyn M, De Vuyst H, et al. European guidelines for quality assurance in cervical cancer screening. Summary of the supplements on HPV screening and vaccination. Papillomavirus Res. 2015;1:22–31.
[2] Chapter in a book: Therapy and vaccination. In: IARC monographs on the evaluation of carcinogenic risks to humans, volume 90: human papillomaviruses. Lyon: IARC Press; 2007. p. 156–176.
[3] Arbyn M, Anttila A, Jordan J, et al., editors. European guidelines for quality assurance in cervical cancer screening. Brussels: Office for official publications of the European Communities; 2008.
[4] Bjørge T, Skare GB, Bjørge L, et al. Adverse pregnancy outcomes after treatment for cervical intraepithelial neoplasia. Obstet Gynecol. 2016;128:1265–1273.
[5] Pukkala E, Engholm G, Højsgaard Schmidt LK, et al. Nordic cancer registries – an overview of their procedures and data comparability. Acta Oncol. 2017;11:1–16.
[6] Enerly E, Bray F, Mellem C, et al. Quality assessment of the registration of vulvar and vaginal premalignant lesions at the Cancer Registry of Norway. Acta Oncol. 2012;51:45–50.
[7] GB, Skare S, Lönnberg T, Bjørge, et al. Livmorhalsprogrammet. Årsrapport 2015. Oslo: Cancer Registry of Norway; 2016. [In Norwegian]
[8] Mallin K, Palis BE, Watroba N, et al. Completeness of American cancer registry treatment data: implications for quality of care research. J Am Coll Surg. 2013;216:428–437.
[9] Norwegian SNOMED Coding System. [cited 2018 Jan 29]. Available from: https://www.kreftregisteret.no/Registrene/Innrapportering/Rapportering-av-patologiinformasjon/
[10] Fritz A, Percy C, Jack A, et al., editors. International classification of diseases for oncology. 3rd ed., First Revision. Geneva: World Health Organization; 2013.
[11] Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
[12] Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
[13] Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8:206–213.
[14] Kyrgiou M, Athanasiou A, Paraskevaidi M, et al. Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis. BMJ 2016;354:i3633.
[15] Khan MJ, Smith-McCune KK. Treatment of cervical precancers: back to basics. Obstet Gynecol. 2014;123:1339–1343.
[16] Wright TC, Jr Massad LS, Dunton CJ, et al. 2006 consensus guidelines for the management of women with cervical intraepithelial neoplasia or adenocarcinoma in situ. Am J Obstet Gynecol. 2007;197:340–345.

[17] Moscicki AB, Ma Y, Wibbelsman C, et al. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. Obstet Gynecol. 2010;116:1373–1380.
[18] Ho GY, Einstein MH, Romney SL, et al. Risk factors for persistent cervical intraepithelial neoplasia grades 1 and 2: managed by watchful waiting. J Low Genit Tract Dis. 2011;15:268–275.
[19] McAllum B, Sykes PH, Sadler L, et al. Is the treatment of CIN 2 always necessary in women under 25 years old? Am J Obstet Gynecol. 2011;205:478.e1–477.
[20] Discacciati MG, de Souza CA, d'Otavianno MG, et al. Outcome of expectant management of cervical intraepithelial neoplasia grade 2 in women followed for 12 months. Eur J Obstet Gynecol Reprod Biol. 2011;155:204–208.
[21] Piris S, Bravo V, Alvarez C, et al. Natural history of histologically moderate cervical dysplasia in adolescent and young women. Onco Targets Ther. 2014;7:2101–2106.
[22] Waxman AG, Chelmow D, Darragh TM, et al. Revised terminology for cervical histopathology and its implications for management of high-grade squamous intraepithelial lesions of the cervix. Obstet Gynecol. 2012;120:1465–1471.
[23] Stoler MH, Ronnett BM, Joste NE, et al. The Interpretive variability of cervical biopsies and its relationship to HPV status. Am J Surg Pathol. 2015;39:729–736.
[24] Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009;45:747
[25] Cancer Registry of Norway. Cancer in Norway 2016 – cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway; 2017.
[26] Borch KB, Weiderpass E, Braaten T, et al. Physical activity and risk of endometrial cancer in the Norwegian Women and Cancer (NOWAC) study. Int J Cancer. 2017;140:1809–1818.
[27] Pearson-Stuttard J, Zhou B, Kontis V, et al. Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment. Lancet Diabetes Endocrinol. 2018;6:92–95.
[28] Coebergh JW, van den Hurk C, Rosso S, et al. EUROCOURSE lessons learned from and for population-based cancer registries in Europe and their programme owners: improving performance by research programming for public health and clinical evaluation. Eur J Cancer. 2015;51:997.
[29] Batra S, Sachdeva S. Organizing standardized electronic healthcare records data for mining. Health Policy Technol. 2016;5:226–242.
[30] Mooney SJ, Pejaver V. Big data in public health: terminology, machine learning, and privacy. Annu Rev Public Health. 2017.
[31] Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Wiens Clin Infect Dis. 2018;66:149–153.
[32] van Walraven C, Bennett C, Forster AJ. Administrative database research infrequently used validated diagnostic or procedural codes. J Clin Epidemiol. 2011;64:1054–1059.
[33] Sund R. Quality of the Finnish hospital discharge register: a systematic review. Scand J Public Health. 2012;40:505.
[34] Bakken IJ, Gystad SO, Christensen OO, et al. Comparison of data from the Norwegian Patient Register and the Cancer Registry of Norway. Tidsskr nor Laegeforen. 2012;132:1336–1340. [In Norwegian]
[35] Beatty JD, Adachi M, Bonham C, et al. Utilization of cancer registry data for monitoring quality of care. Am J Surg. 2011;201:645–649.
[36] Caldarella A, Amunni G, Angiolini C, et al. Feasibility of evaluating quality cancer care using registry data and electronic health records: a population-based study. Int J Qual Health Care. 2012;24:411–418.
[37] Gurney J, Sarfati D, Dennett E, et al. The completeness of cancer treatment data on the National Health Collections. N Z Med J. 2013;126:69–74.
[38] Siesling S, Louwman WJ, Kwast A, et al. Uses of cancer registries for public health and clinical research in Europe: results of the European Network of Cancer Registries survey among 161 population-based cancer registries during 2010–2012. Eur J Cancer. 2015;51:1039–1049.