HHS Public Access

Author manuscript

Br J Dermatol. Author manuscript; available in PMC 2018 May 01.

Published in final edited form as:

Br J Dermatol. 2017 May ; 176(5): 1363–1365. doi:10.1111/bjd.14957.

Validity of diagnosis codes for identifying cutaneous squamous cell carcinoma in The Health Improvement Network

Z.C. Chiesa Fuxench[1], **A.B. Troxel**[2], and **J.M. Gelfand**[1,2]

[1]Department of Dermatology, University of Pennsylvania Perelman School of Medicine, Philadelphia

[2]Departments of Epidemiology and Biostatistics, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania Perelman School of Medicine, Philadelphia

Basal cell carcinoma (BCC) and cutaneous squamous cell carcinoma (cSCC), collectively referred to as “non-melanoma skin cancer (NMSC)”, are the most common cancers worldwide.[1,2] Despite their considerable public health burden, there are substantial gaps in the epidemiological study of these malignancies as they are not commonly included in cancer registries.[3,4] The Health Improvement Network (THIN) is a medical records database that contains anonymous medical record information from practices across the United Kingdom (UK).[5–9] Prior research has demonstrated that diagnostic codes accurately identify patients with BCC in THIN.[8] However, the positive predictive value (PPV) of diagnostic codes for cSCC has not been investigated.

We performed a cross-sectional study to determine the PPV of diagnosis codes for identifying cSCC in THIN using the general practitioner’s (GP) confirmation as the reference standard. We designed a questionnaire to collect information relevant to the diagnosis of cSCC (SF1) and reviewed the patient’s medical records for additional information. Patients identified for study inclusion were between the ages of 18–89 and had at least one diagnostic code for SCC recorded in THIN between July 1st 2012 and December 31st 2012. Our sample size consisted of 100 randomly selected patients. We predicted, *a priori*, that 90% of patients who received a diagnosis of SCC in THIN do have cSCC, therefore if we used our selected algorithm, then 95% of the time we would obtain a PPV that lies between 82% and 95%. Descriptive statistics were used to compare our groups using the student’s t-test or its non-parametric equivalent for continuous variables and the χ^2 test for categorical variables. The PPV (95%CI) was calculated with the GP’s confirmed diagnosis as the reference standard. We also determined the sensitivity and specificity for algorithm ii (having more than one SCC code). For this algorithm, the presence or absence of an additional SCC code was considered the “test”. Finally, we analyzed the PPV of

Correspondence: Zelma C. Chiesa Fuxench, MD MSCE, Department of Dermatology, University of Pennsylvania Perelman School of Medicine, 3400 Spruce Street, 1107 Dulles, Philadelphia, PA 19104, zelma.chiesafuxench@uphs.upenn.edu.

Institution at which research was conducted: University of Pennsylvania Perelman School of Medicine

Conflict of Interest/Financial Disclosure: Dr. Gelfand served as a consultant, for AbbVie, Amgen Inc., Coherus, Celgene Corp, Eli Lilly, Merck, Janssen Biologics (formerly Centocor), Regeneron Novartis Corp., Leo Pharmaceuticals, Endo and Pfizer Inc., receiving honoraria; had grants or has pending grants from AbbVie, Janssen, Regeneron, Novartis Corp., and Pfizer Corp. Dr. Chiesa Fuxench and Dr. Troxel have no financial interests to disclose.

having either a code for SCC of the Skin (B338.00) or SCC NOS (BB2A.00) as these codes were found to be among the most frequently used by GPs in THIN. This study was developed in accordance to the STROBE guidelines and was granted exemption from review by the institutional review board of the University of Pennsylvania Perelman School of Medicine and the Scientific Review Committee of CSD Medical Research, UK.

Of 100 questionnaires initially sent out, 85 were returned within a 6-month period. Four questionnaires (4.71%) were returned incorrectly filled and despite further efforts for verification no additional information was obtained. We found no statistically significant differences in the baseline characteristics of patients with returned and not returned questionnaires (data not shown). Eighty-one questionnaires were included in our analyses (Table 1). The most commonly used codes in cases confirmed as cSCC were BB2a.00 ([M] SCC non-otherwise specified) and B338.00 (SCC of skin) in 34/67 (50.75%) and 25/67 (37.31%) of cases, respectively. Of patients confirmed as cSCC, 42% had more than one code for SCC recorded in THIN whereas only 14% of cases confirmed as not cSCC had more than one SCC code. The PPV for any SCC code was 82.72%. The PPV for addition of a second code for SCC was 93.33%, however the sensitivity (45%) and specificity (55%) of this algorithm was low. Examination of the PPV for codes SCC NOS (BB2A.00) or SCC of the skin (B338.00) combined yielded a PPV of 84.29% (Table 2).

Meal et al. showed that codes used for BCC in THIN had a PPV of 93%; yet, no data were presented on the validity of codes for identifying cSCC.[8] In this study we demonstrate that certain diagnostic codes used in THIN to code SCC can reliably identify patients with cSCC with a PPV of 83%. We also found that addition of a second SCC code recorded on a different date resulted in an increase in the PPV although the sensitivity and specificity of this algorithm were low. As with all studies, there are important limitations to consider. Our study only included cases from practices that had previously agreed to participate in validation studies potentially limiting the generalizability of our results, as these practices may be different from practices that do not participate in validation studies. We were unable to examine the sensitivity and specificity of the primary algorithm since we could not verify the status of patients who might have had a diagnosis of cSCC but did not have a diagnostic code recorded in THIN). Lastly, although our PPV is within generally accepted ranges for validity, a small percentage of internal SCCs were misclassified as cSCC. Because these internal cancers are associated with high morbidity and mortality caution needs to be exercised when interpreting cSCC outcomes identified via the coding algorithm herein. Despite these limitations, our findings demonstrate that there is potential for THIN to be useful in improving the epidemiological study of cSCC.

Acknowledgments

Funding/Support: This study was supported by a Training Grant 5T32GM075766-09 from the National Institutes of Health (Chiesa Fuxench), NIAMS K24-AR064310 (JMG), and an institutional grant from the Center for Clinical Epidemiology and Biostatistics of the University of Pennsylvania Perelman School of Medicine(Chiesa Fuxench). No study sponsors were involved in the study design, collection, analysis and interpretation of the data, or in drafting of the manuscript.

References

1. Cakir BO, Adamson P, Cingi C. Epidemiology and economic burden of nonmelanoma skin cancer. Facial Plasc Surg Clin North Am. 2012; 20:419–422.
2. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the US Population, 2012. JAMA dermatol. 2015; 151:1081–1086. [PubMed: 25928283]
3. Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. Br J Dermatol. 2012; 166:1069–1080. [PubMed: 22251204]
4. Marcil I, Stern RS. Risk of developing a subsequent nonmelanoma skin cancer in patients with a history of nonmelanoma skin cancer: a critical review of the literature and meta-analysis. Arch Dermatol. 2000; 136:1524–1530. [PubMed: 11115165]
5. Dregan A, Moller H, Murray-Thomas T, et al. Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. Cancer Epidemiol. 2012; 36:425–429. [PubMed: 22727737]
6. Haynes K, Forde KA, Schinnar R, Wong P, et al. Cancer incidence in The Health Improvement Network. Pharmacoepidemiol Drug Saf. 2009; 18:730–736. [PubMed: 19479713]
7. Seminara NM, Abuabara K, Shin DB, et al. Validity of The Health Improvement Network (THIN) for the study of psoriasis. Br J Dermatol. 2011; 164:602–609. [PubMed: 21073449]
8. Meal A, Leonardi-Bee J, Smith C, et al. Validation of THIN data for non-melanoma skin cancer. Qual Primary Care. 16:49–52.
9. Ogdie A, Alehashemi S, Love TJ, et al. Validity of psoriatic arthritis and capture of disease modifying antirheumatic drugs in the health improvement network. Pharmacoepidemiol Drug Saf. 2014; 23:918–922. [PubMed: 25044030]

Table 1

Characteristics of study patients for whom questionnaires were received

	Confirmed cSCC	Confirmed Not cSCC	*P*-value
N (%)	**67 (82.72)**	**14 (17.28)**	
Age, mean (SD)	73.81 (11.69)	70.78 (10.40)	0.08
Sex, *N (%)*			
Male	38 (56.72)	9 (64.29)	0.77
Female	29 (43.28)	5 (35.71)	
SCC codes, *N (%)*			0.11
B338.00 (SCC of skin)	25 (37.31)	5 (35.71)	
BB2A.00 ([M] SCC NOS)	34 (50.75)	6 (42.86)	
B33z.00 (Malignant neoplasm of skin NOS)	1 (1.49)	2 (14.28)	
BB2L.00 ([M]Bowen's disease)	6 (8.96)	0 (0.00)	–
BB29.12 ([M]Intraepidermal carcinoma NOS	1 (1.49)	1 (7.14)	
Number of SCC codes, *N (%)*			
>1 code	28 (41.79)	2 (14.29)	0.07
1 code	39 (58.21)	12 (85.71)	
Location of specific lesion, *N(%)* [a]			
Head/Neck	33 (49.25)	N/A	
Chest/Back	4 (5.97)	N/A	
Upper/Lower Extremity	22 (33.84)	N/A	
Physician diagnosis, *N(%)* [b]			
General Practitioner	14 (20.90)	N/A	
Dermatologist	29 (43.28)	N/A	
Plastic or general surgeon	10 (14.93)	N/A	
Confirmed with skin biopsy, *N(%)*	59 (88.06)	N/A	
Referred to specialty clinic, *N(%)* [c]	34 (50.74)	N/A	
Reason for misclassification as SCC of the skin, *N(%)*		2 (14.29)	
Coding error		1 (7.14)	
Unknown			
Patient had a different skin diagnosis, *N(%)*		6 (42.86)	
Basal cell carcinoma		4 (66.67)	
Intraepidermal carcinoma, precancerous		1 (16.67)	
Scar revision		1 (16.67)	
Patient had a diagnosis of SCC but NOT on the skin, *N(%)* [d]		5 (35.71)	

[a] 3 (4.84%) additional cases of SCC localized to other areas: right groin, anal, tongue

[b] 9 (14.52%) additional cases of SCC diagnosed by: Oral surgery(1), ENT(2), head and neck oncology(1), ophthalmology (1), maxillofacial(4)

[c] This includes referral to dermatology clinic or other community skin cancer clinic

[d] SCC classified as not localized to the skin where located in the rectum/anus (1/5: 20%); larynx (1/5; 20%); vulva (1/5; 2%) and tonsils (2/5; 40%)

Table 2

Positive predictive value of SCC codes in THIN for identifying cSCC

	No. Cases	No. True Positives	PPV (95%CI)	Sn	Sp
i. Any SCC code	81	67	82.72 (72.70, 90.22)	N/A	N/A
ii. More than one SCC code	30	28	93.33 (77.93, 99.18)	45.16	54.84
Sensitivity Analyses					
B338.00 or BB2A.00[a]	70	59	84.29 (73.62, 91.89)	N/A	N/A

[a]Squamous Cell Carcinoma of the Skin (B338.00) and Squamous Cell Carcinoma NOS (BB2A.00)