Published in final edited form as:

Pharmacoepidemiol Drug Saf. 2016 April ; 25(4): 399–404. doi:10.1002/pds.3952.

Validity of Ankylosing Spondylitis Diagnoses in The Health Improvement Network

Maureen Dubreuil^{1,2}, Christine Peloquin³, Yuqing Zhang³, Hyon K. Choi⁴, Robert D. Inman⁵, and Tuhina Neogi³

¹Boston University School of Medicine, Section of Rheumatology; Boston, MA, USA

²VA Boston Healthcare System; Boston, MA, USA

³Boston University School of Medicine, Section of Clinical Epidemiology; Boston, MA, USA

⁴Massachusetts General Hospital, Section of Rheumatology; Boston, MA, USA

⁵University of Toronto, Toronto Western Hospital; Toronto, ON, Canada

Keywords

Ankylosing spondylitis; electronic medical record

Background

Ankylosing spondylitis (AS) is a chronic progressive inflammatory arthritis of the axial skeleton, and is in the family of diseases encompassed by the term “axial spondyloarthritis” (axSpA). Classification criteria for axSpA allow identification of affected persons for research purposes, and include clinical features, such as age of symptoms onset, back pain features and extra-spinal manifestations, as well as laboratory and radiologic features.¹ Retrospective application of these criteria within a pre-existing electronic health record (EHR) database is challenging because such data may exist only as text within a detailed clinical note, or may be missing entirely.²³ Nonetheless, large EHR databases offer important opportunities for epidemiologic research on AS because a large population study may confirm the results of small studies that are prone to random error, provide the power to estimate effects more precisely than smaller studies, and examine risk factors or events that occur too infrequently to be captured in small studies. The longitudinal nature of EHR databases also allows investigation of associations with a long latency from exposure to outcome, time-varying exposures, and secular trends. The strengths of large EHR databases make them particularly well suited for pharmacoepidemiologic research in AS, a relatively uncommon but chronic disease for which multiple treatments may be used over time, and in which many features and comorbidities take years to develop. Therefore, we sought to

Corresponding Author: Maureen Dubreuil, MD, 650 Albany St, Suite 200, Boston, MA 02118, Phone: 617-638-4394, Fax: 617-638-5239.

Conflict of Interest: None of the authors report financial conflicts of interest related to this work.

Prior postings and presentations: This work has not previously been presented or published.

validate diagnoses of AS in The Health Improvement Network (THIN), a large EHR database from the United Kingdom.

Methods

The THIN database contains electronic health records for over 10 million persons in the UK, with data entered by about 580 general practices as of 2013.⁴ The database includes clinical information obtained or determined by the GP, such as diagnoses, details on medication use and prescriptions, hospital admissions and discharges, surgical procedures, laboratory and radiographic testing, and results of consultations.⁵ Diagnoses are recorded using the Read classification, and drugs are codified according to the Multilex system. Though many data exist as structured elements that are easily accessible to researchers, researchers may not readily have access to data generated by consultants outside of the EHR, as such data may exist only in the context of a detailed clinical report, which is costly to obtain and labor-intensive to review.³ Thus, to determine the accuracy of AS diagnostic codes, we administered a questionnaire to the GPs caring for patients with one or more AS diagnoses (Read code N100.00, “Ankylosing spondylitis”) in THIN.⁶⁷ This method was selected for AS diagnosis validation because of the potential for diagnoses to be miscoded in GP records, such as in the event that a tentative diagnosis was recorded using a code that failed to capture the diagnostic uncertainty.

Eligibility criteria

We established eligibility criteria similar to those typically used in epidemiologic studies using an EHR database; age 18–59 years, at least one year enrollment in THIN both before and after the first AS diagnosis, and at least one GP visit after AS diagnosis.⁸⁹ We also required included subjects to be followed in a practice that consented to research questionnaires.⁶ We included THIN data from the years 2000 to 2013. Due to financial restrictions, we were able to query GPs for 100 persons with AS diagnoses, and the included subjects were randomly selected from the total eligible population.

Questionnaire development and administration

Our priority in questionnaire development was assessment of the positive predictive value (PPV) of AS diagnostic codes and other AS case ascertainment algorithms. PPV was emphasized because in epidemiologic studies of uncommon diseases, it is necessary to identify a population that is truly diseased. If the “diseased” cohort contains many persons without the disease of interest (misclassification of exposure), event rates in the “diseased” cohort would occur at rates more similar to the non-diseased population, and study results would likely be biased toward the null. High PPV is more important in studies of AS than high NPV because AS is an uncommon disease,¹⁰ and it is therefore unlikely that a person without an AS diagnosis would have AS and be selected for a comparator population. Our secondary goal was to determine algorithm sensitivity because high sensitivity would allow identification of the greatest number of AS cases possible, and thereby improve the ability to capture infrequent events and detect small effect sizes in epidemiologic studies.

The questionnaire was developed with AS expert input (RI) to include confirmation of the AS diagnosis by the GP. We also asked the GP to review the subjects record and report if the diagnosis was confirmed by a rheumatologist, age of AS onset, use of medications, features of inflammatory back pain, and each of the individual elements of the Assessment of SpondyloArthritis International Society (ASAS) classification criteria for axial spondyloarthritis, including arthritis, enthesitis, uveitis, psoriasis, radiographic findings inflammatory markers and HLA-B27 status.¹ The questionnaire was distributed in July 2014, and those returned by March 2015 were included in the analysis. In the event of non-response after 30 days, a reminder was sent, and the same was done at 60 days after the initial questionnaire distribution.

Data analysis

We calculated PPV and sensitivity with 95% confidence intervals for each of the AS identification algorithms. PPV was calculated as the proportion of subjects whose GP confirmed an AS diagnosis (our diagnostic gold standard), among those who fulfilled the algorithm. Sensitivity was calculated as the proportion of GP-confirmed AS cases who met the algorithm, within the population we sampled (at least one AS diagnosis); this does not represent overall sensitivity in the database, as we did not sample subjects without an AS code. Other case identification algorithms included 1 or more AS codes, and combinations of AS codes with other clinical features. In particular, we assessed an algorithm requiring 2 AS diagnoses separated by ≥ 7 days because the separation in time suggests that the subject had at least 2 separate encounters with the practice for the same condition (whereas diagnoses recorded within a few days of each other are more likely to result from a single encounter). Although GP diagnosis of AS, as documented on the questionnaire, was considered the gold standard to which the case identification algorithms were compared, we also considered alternative definitions of the gold standard, including a rheumatologist-confirmed AS diagnosis, and fulfilment of ASAS classification criteria for axial spondyloarthritis. We reported descriptive statistics for the subjects in whom the AS diagnosis was confirmed by the GP.

This study was approved by the UK Medical Research Scientific Review Committee (Protocol 13047), and judged exempt from review by the Boston University Institutional Review Board (H-32735)

Results

The inclusion criteria were satisfied by 606 subjects, of which 100 subjects were randomly selected for questionnaire administration (Figure 1). Following questionnaire distribution to the GPs, one initially eligible subject disenrolled from a practice that had consented to research participation, and was subsequently excluded from this study. Of the 99 remaining questionnaires, 77 were returned within 90 days, and 85 were returned by the end of the study period.

AS was confirmed by the GP in 61 of the 85 subjects, and the clinical characteristics the confirmed AS subjects are summarized below (Table 1). The majority were male (73.7%) and mean age at AS diagnosis was 38 (± 9) years. The GP reported at least one feature of

inflammatory back pain (IBP) in half of confirmed AS cases (54.1%). However, features of IBP were unknown by the GP for 37.7% of subjects. Although a good response to NSAIDs was documented by the GP in more than 70% of confirmed AS cases, documented use of a DMARD or biologic drug was uncommon (13.1% and 36.1%, respectively). Peripheral arthritis was common (41.0%), as was psoriasis (19.7%). Laboratory and radiographic features of spondyloarthritis were also common (24.6–62.3%).

Considering the GP's clinical impression as the gold standard, an AS diagnostic code had an overall PPV of 71.8% (95% CI 61.0–81.0; Table 2). PPV was greatest for the algorithms requiring 2 AS diagnoses separated in time by 7 days (PPV 88.6%), and an AS diagnosis in combination with a DMARD or biologic prescription (PPV 85.7%). PPV was 71–75% for the other algorithms tested. Sensitivity was 63.9% for the algorithm requiring 2 diagnoses separated by a week, however sensitivity was low when a prescription for DMARD/biologic was required (29.5%). When the rheumatologist's diagnosis was considered the gold standard to which each AS case identification algorithm was compared, PPVs were slightly lower (PPV 68% for the 1 AS diagnosis algorithm and 84% for the algorithm requiring 2 AS diagnosis separated by 7 days; Table 3). If ASAS criteria for axial spondyloarthritis was considered the gold standard for diagnosis, the PPVs were 72% and 77% for the 1 diagnosis and 2 diagnosis algorithms, respectively.

Discussion

This study of the accuracy of diagnostic codes and other algorithms to identify AS cases in THIN found that, overall, a diagnostic code for AS was associated with a PPV of 71.8% relative to the GP's clinical impression as the gold standard. Higher PPV of 89%, was achieved by requiring 2 AS diagnoses separated by 7 days, or a diagnosis in combination with a DMARD or biologic drug prescription. We recommend application of the two diagnosis algorithm, which has the highest PPV, for use in epidemiologic studies of AS in which accurate identification of AS cases is the priority. In this database of GP records, there was insufficient data to allow AS case identification on the basis of other clinical data, such as elements of the ASAS Classification criteria.

There is little published data on the validation of AS diagnoses in large datasets. In the Swedish National Patient Register, which includes hospital diagnoses but not outpatient records, ICD codes for AS had an 89% PPV relative to any set of SpA criteria.¹¹ ICD codes for AS were also assessed among patients seen in a United States VA rheumatology clinic; PPV for 2 diagnostic codes in this setting was 100%.¹² Our study in a GP-based EHR found a lower PPV, possibly reflecting that GPs occasionally record a tentative AS diagnosis in the process of working up a patient's condition then later revise the diagnosis when additional clinical information becomes available. Our findings are largely similar to studies on the validity of psoriatic arthritis (PsA) diagnoses in large datasets, including THIN.⁶ Notably, our results of 88.6% PPV using a fairly simple algorithm are close to the PPV achieved using a much more labor-intensive process, natural language processing (NLP) in a large dataset to identify PsA (PPV 90%).¹³ Also similar to our findings, the additional requirement for a DMARD or biologic prescription resulted in decreased sensitivity (27–65%) in previous studies of PsA and AS.^{6,12} This pattern suggests that the majority of AS

patients did not receive therapy with a DMARD or biologic, but may also reflect that prescriptions from specialists are inconsistently documented in GP records and that EHRs currently lack interoperability between specialists and GPs.¹⁴

In addition to the estimates of PPV and sensitivity presented, this work highlights several important aspects of research using large EHR databases. Firstly, GP awareness and documentation of inflammatory back pain features is not universal. For many rheumatologists, the distinction between inflammatory and mechanical features of back pain is the first step in identifying persons with axial spondyloarthritis. In this study, more than 37% of GPs did not know if their AS patient had inflammatory back pain. Thus, there is an opportunity to enhance GP education regarding the importance of identifying IBP. Secondly, documentation of DMARDs and biologic medication use was incomplete in THIN. Although GPs reported DMARD and biologic use in 13% and 36% of AS patients through the questionnaire, by review of the electronic prescription records in THIN, 29.5% of AS patients had been prescribed DMARDs, and biologic use had been documented in zero patients. Thus, DMARD use may be underrecognized by GPs, and biologic use underdocumented, possibly because these medications are prescribed by specialists who work outside of the GP-based EHR. Incomplete recognition and recording of DMARD and biologic use not only poses a challenge to researchers using these databases, but raises questions regarding the GP's role in the clinical assessment of medication side effects.

There are several potential mechanisms for AS case identification that we were not able to assess in this study. For example, it is likely that the presence of 3, 4 or 5 AS diagnostic codes per subject is likely associated with a greater PPV, however, 3 codes occurred too infrequently to allow precise estimation of PPV. In this study, we were also unable to assess true sensitivity or specificity of AS diagnostic codes, as we could not readily identify a population for whom the reference standard (GP diagnosis) was present without use of AS diagnostic codes, and we lacked the financial resources to confirm the absence of AS in the general population. Also, as discussed above, although ASAS classification criteria are ideal for identification of persons with axial SpA, application of these criteria in a GP-derived EHR is not feasible.

This study has several strengths, including the high questionnaire response rate (86%), our application of inclusion criteria that are commonly used for epidemiologic studies, and random sampling of subjects from the eligible population. The results of this work are generalizable to other studies using THIN, and likely to the Clinical Practice Research Datalink (CPRD), as THIN and CPRD share approximately half of the practices.⁴ The lessons learned from this study of AS may also apply to studies of other uncommon diseases that are primarily diagnosed or treated by medical subspecialists, in that application of additional case identification criteria beyond a single diagnostic code may improve PPV.

In conclusion, this study found that identification of AS cases by the presence of 1 AS diagnostic codes had low PPV in identifying AS cases. An algorithm requiring 2 or more diagnoses separated by 7 days had the highest PPV and could be used to optimize accurate AS case identification in epidemiologic studies.

Acknowledgments

Sponsors/Funding: The Arthritis Foundation (MD), the Department of Veterans Affairs (MD) and NIH AR47785.

References

1. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis.* 2009; 68:777–783. [PubMed: 19297344]
2. Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. *BMJ.* 1991; 302(6779):766–768. [PubMed: 2021768]
3. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet.* 2012; 13(6):395–405. [PubMed: 22549152]
4. Ogdie, A.; Langan, S.; Parkinson, J.; Dattani, H.; Kostev, K.; Gelfand, J. Medical Record Databases. In: Strom, BKS.; Hennessy, S., editors. *Pharmacoepidemiology*. Oxford, UK: Wiley-Blackwell; 2012. p. 224–243.
5. IMS Health. THIN: Data Content. London, UK: IMS Health Incorporated; 2015.
6. Ogdie A, Alehashemi S, Love TJ, Jiang Y, Haynes K, Hennessy S, et al. Validity of psoriatic arthritis and capture of disease modifying antirheumatic drugs in the health improvement network. *Pharmacoepidemiol Drug Saf.* 2014; 23(9):918–922. [PubMed: 25044030]
7. Kang EM, Pinheiro SP, Hammad TA, Abou-Ali A. Evaluating the validity of clinical codes to identify cataract and glaucoma in the UK Clinical Practice Research Datalink. *Pharmacoepidemiol Drug Saf.* 2015; 24(1):38–44. [PubMed: 25425327]
8. Scowcroft AC, Lee S, Mant J. Thromboprophylaxis of elderly patients with AF in the UK: an analysis using the General Practice Research Database (GPRD) 2000–2009. *Heart.* 2013; 99(2): 127–132. [PubMed: 23086966]
9. Lapi F, Azoulay L, Yin H, Nessim SJ, Suissa S. Concurrent use of diuretics, angiotensin converting enzyme inhibitors, and angiotensin receptor blockers with non-steroidal anti-inflammatory drugs and risk of acute kidney injury: nested case-control study. *BMJ.* 2013 Jan.8(346)
10. Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of ankylosing spondylitis. *Rheumatology (Oxford).* 2014; 53(4):650–657. [PubMed: 24324212]
11. Lindström U, Exarchou S, Sigurdardottir V, Sundström B, Askling J, Eriksson JK, et al. Validity of ankylosing spondylitis and undifferentiated spondyloarthritis diagnoses in the Swedish National Patient Register. *Scand J Rheumatol.* 2015; 23:1–8.
12. Singh JA, Holmgren A, Krug H, Noorbaloochi S. Accuracy of the diagnoses of spondylarthritides in veterans affairs medical center databases. *Arthritis Rheum.* 2007; 57(4):648–655. [PubMed: 17471541]
13. Love TJ, Cai T, Karlson EW. Validation of psoriatic arthritis diagnoses in electronic medical records using natural language processing. *Semin Arthritis Rheum.* 2011; 40(5):413–420. [PubMed: 20701955]
14. Wachter RM. Why Health Care Tech Is Still So Bad. *New York Times.* 2015

Key points

- An ankylosing spondylitis case identification algorithm of two diagnostic codes separated by at least 7 days provided the highest positive predictive value and may be used in ankylosing spondylitis research.
- Biologic use in ankylosing spondylitis is incompletely documented by general practitioners in this electronic medical record database.

Figure 1.
Subject selection and response to questionnaires

Table 1

Clinical, laboratory and radiographic features among 61 subjects with a GP-confirmed diagnosis of AS.

Demographics	N (%)
Male	45 (73.7)
Age at first AS diagnosis, mean (SD)	37.6 (9.2)
AS features on database review	
1 AS code	22 (36.0)
2 or more AS codes	39 (64.0)
NSAID prescription	60 (98.4)
DMARD prescription	18 (29.5)
Biologic prescription	0
DMARD or biologic prescription	18
NSAID <u>and</u> DMARD/biologic prescription	18
Years follow up after AS diagnosis, mean (SD)	8.0 (3.5)
Features of inflammatory back pain (IBP) *	
Insidious onset	26 (42.6)
Improvement with exercise	15 (24.6)
Lack of improvement with rest	8 (13.1)
Pain at night	10 (16.4)
None of the above	5 (8.2)
Unknown	23 (37.7)
Use of medications by GP questionnaire	
Glucocorticoids	48 (78.7)
DMARDs	8 (13.1)
Biologics	22 (36.1)
Other Clinical Features *	
Arthritis	25 (41.0)
Enthesitis	6 (9.8)
Dactylitis	3 (4.9)
Psoriasis	12 (19.7)
Inflammatory bowel disease	4 (6.6)
Good response to NSAIDs	43 (70.5)
Family History of spondyloarthritis	14 (23.0)
Radiographic/Laboratory Features *	
Sacroiliitis by X-ray	34 (55.7)
Sacroiliitis by MRI	15 (24.6)
Elevated ESR or CRP	38 (62.3)
Positive HLA-B27	34 (55.7)

* Ascertained by GP questionnaire.

NSAID: nonsteroidal anti-inflammatory drug; DMARD: disease-modifying antirheumatic drug; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein

Table 2

Algorithms for Ankylosing Spondylitis case identification in The Health Improvement Network

Algorithm	Total satisfying algorithm (N)	Confirmed AS (N)	PPV, % (95% CI)	Sensitivity*
1 AS Code	85	61	71.8 (61.0–81.0)	N/A
2 AS Codes 7 days apart	44	39	88.6 (75.4–96.2)	63.9 (50.1–75.8)
1 AS Code <i>and</i> absence of OA code	77	58	75.3 (76.0–97.3)	95.1 (89.5–100)
1 AS Code <i>and</i> absence of RA code	80	58	72.5 (61.0–84.0)	95.1 (89.5–100)
1 AS Code <i>and</i> DMARD or Biologic	21	15	85.7 (63.7–97.0)	29.5 (18.5–42.6)
1 AS Code <i>and</i> NSAID	84	60	71.4 (60.0–82.9)	98.4 (95.1–100)

AS = ankylosing spondylitis, DMARD = Disease Modifying Antirheumatic Drug, NSAID =nonsteroidal anti-inflammatory drug, OA = osteoarthritis, RA = rheumatoid arthritis, PPV= positive predictive value.

* Sensitivity of the algorithm relative to a confirmed AS diagnosis of AS according to GP questionnaire, among patients with at least 1 code for AS.

Table 3

Positive predictive value and sensitivity estimates for selected AS case identification algorithms relative to alternative gold standard definitions of AS diagnosis

Algorithm	Gold Standard Definitions							
	GP diagnosis of AS		Rheumatologist diagnosis of AS		ASAS Criteria		Rheumatologist OR ASAS Criteria	
	PPV (95% CI)	Sensitivity (95% CI)*	PPV (95% CI)	Sensitivity (95% CI)*	PPV (95% CI)	Sensitivity (95% CI)*	PPV (95% CI)	Sensitivity (95% CI)*
1 AS Code	72% (61–81)	N/A	68% (57–78)	N/A	72% (60–83)	N/A	69% (58–79)	N/A
2 AS Codes 7 days apart	89% (75–96)	64% (50–76)	84% (70–93)	64% (50–76)	77% (61–89)	67% (51–80)	86% (73–95)	64% (51–76)
1 AS Code and DMARD or biologic	86% (64–97)	30% (19–43)	86% (64–97)	31% (20–45)	72% (47–90)	29% (16–44)	86% (64–97)	31% (19–44)

AS = ankylosing spondylitis, DMARD = Disease Modifying Antirheumatic Drug, ASAS=Assessment of SpondyloArthritis International Society, PPV= positive predictive value.

* Sensitivity of the algorithm relative to the gold standard definition, among patients with at least 1 code for AS.