

## Review

# The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank

Frederick Wolfe<sup>1</sup> and Kaleb Michaud<sup>2</sup>

## Abstract

The National Data Bank (NDB) for rheumatic diseases is a patient-based multi-disease, multi-purpose rheumatic disease registry that has been used primarily to study patients with RA, SLE, FM and OA. It enrolls patients from the community, follows up with questionnaires and validates key patient data using medical records. Rheumatologist-written programs make NDB data immediately available to analysts. The NDB has been used to develop and validate diagnostic criteria, develop new questionnaires, describe illness and comorbid disease, assesses disease outcomes and the effect of therapeutic interventions, and measure costs and cost-effectiveness.

**Key words:** NDB, Registry, Data Bank, National Data Bank for rheumatic diseases.

## Introduction

The National Data Bank (NDB) for rheumatic diseases is a rheumatic disease data bank (DB). We prefer the name DB to registry, as registry implies a single purpose activity: a vasculitis registry, a biologics registry and a Type-II diabetes registry. The NDB's scope spans multiple rheumatic diseases, and its activities have multiple purposes that range from statistical teaching and questionnaire development to disease-specific outcome studies.

A DB or DB registry (DBR) represents a research study and data collection effort that differs primarily from other research studies by being longitudinal and ongoing, and almost always without a defined end. In addition, DBRs are actively involved in the collection of data and are responsible for its integrity and quality. DBRs are concerned with four levels of data: personal, physician, external measurement (e.g. laboratory and imaging) and administrative. DBRs may be organized by country or region, social system and disease. A registry addresses a single area, e.g. biologic effectiveness and safety, lupus and RA genetic markers. A DB can be thought of roughly as a collection of registries.

DBRs can be poorly designed and perform poor-quality research. There is nothing about DBRs that makes them

inherently better or less prone to error compared with non-DBR studies. DBRs that are influenced or analysed under the aegis of industry can produce self-serving, biased results.

## The structure and mechanics of the NDB

The NDB is primarily a patient DB in that, in general, it obtains its initial information from patients with rheumatic diseases, and validates this information, when required, from hospital and physician sources and from national death records. The structure, data collection, quality control and research structure of the NDB is shown in Fig. 1.

### Patient source and diagnosis validation

The NDB obtains participants primarily by referral from US and Canadian rheumatologists (Fig. 2). In these cases, the rheumatologist provides the diagnosis. A minority of participants enrol from other sources, including self-referral, after obtaining information from physicians, societies and web sites. In such cases, the NDB obtains diagnostic confirmation from the participants' physicians. Direct enrolment can occur at [www.arthritis-research.org](http://www.arthritis-research.org) or [www.rheummd.org](http://www.rheummd.org).

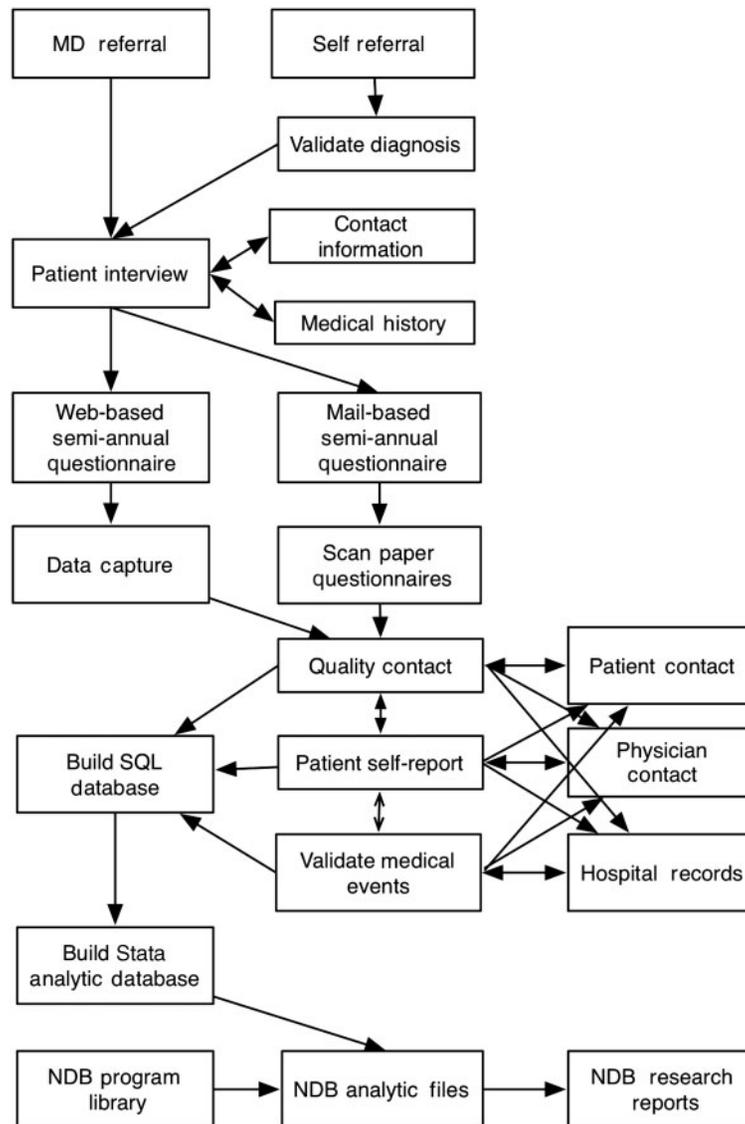
### Initial interview

After preliminary contact, which includes an informed consent, NDB staff contact the patient, usually by telephone, obtain demographic and follow-up information and then obtain a detailed medical history. The history is usually more detailed and relevant than available at a physician's office because it is up-to-date, relevant to the DBR's

<sup>1</sup>National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine, Wichita and <sup>2</sup>National Data Bank for Rheumatic Diseases, University of Nebraska Medical Center, Omaha, NE, USA. Submitted 11 March 2010; accepted 19 April 2010.

Correspondence to: Frederick Wolfe, National Data Bank for Rheumatic Diseases, 1035 N. Emporia, Suite 288, Wichita, KS 67214, USA. E-mail: [fwolfe@arthritis-research.org](mailto:fwolfe@arthritis-research.org)

Fig. 1 Flow diagram of NDB processes.



purpose, and also includes dates of treatments and major medical events.

#### Questionnaire assessments

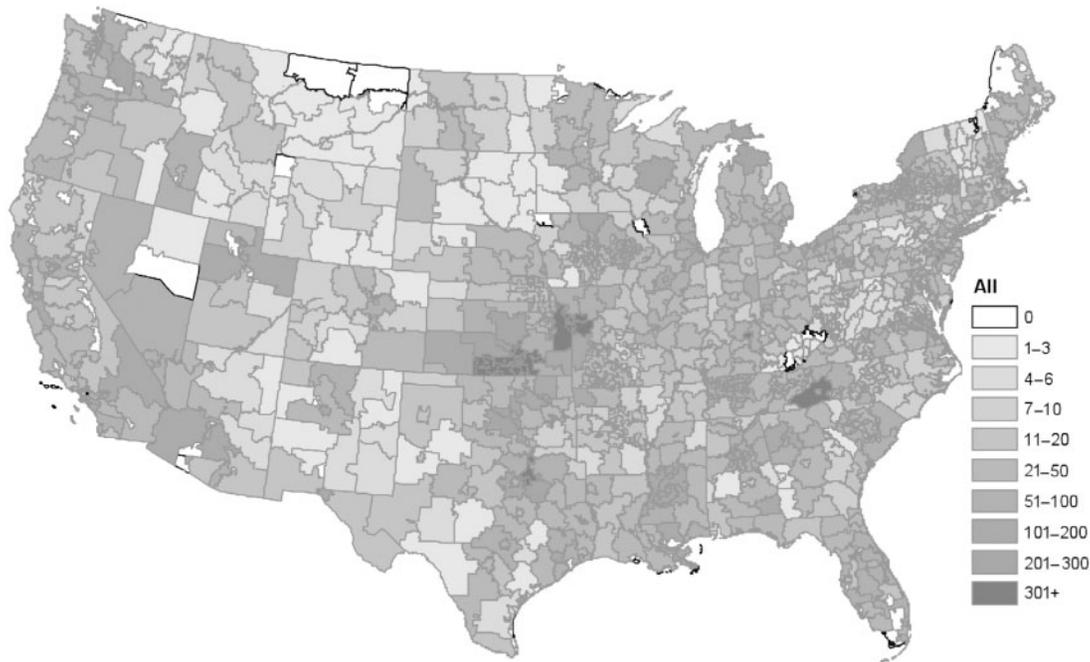
At 6-month intervals (every January and July), NDB participants are surveyed by a postal questionnaire or via the Internet, according to their preferences. A small number of participants who have impairments that prevent the use of these data-collection methods are surveyed by telephone.

The primary semi-annual assessments can be via a comprehensive (28 printed pages or online) questionnaire (<http://www.arthritis-research.org/Documents/Ph58RAFIB.pdf>) or shorter (16 pages) questionnaire. For participants who do not subsequently wish to complete questionnaires, but remain study participants, a brief telephone interview based on the shorter questionnaire is obtained.

#### Quality control and internal validation

Data collection involves Internet surveys, paper questionnaire scanning, patient interviewing, data extraction from medical records, coding and programming. Each of these steps can be a source of data error. Interviewing and coding involves interpretation. The NDB uses coding manuals, ongoing quality control assessments and training sessions to obtain high-quality data. Review of random records and computer programs to trap errors is an integral part of quality control activities. Even so, all errors may not be identified until the data are used in research. Therefore, a key part of the quality control activities of the NDB is exploration of all data in research analyses. The rheumatologist-analyst brings clinical expertise to the quality control process, thereby identifying errors that could be missed by the general staff or statistical quality

Fig. 2 Distribution of US participants in the NDB.



control methods. There is an ongoing checking process that occurs if the analyst finds apparent errors. The analyst communicates the concerns by email or telephone to NDB staff and the staff follow up with data checking, including additional contact with patients and physicians, if required. The NDB staff will then issue a new, dated research DB that contains the corrections.

The NDB also maintains an image DB of all scanned questionnaires. This allows coders and quality control staff to have immediate access to all past questionnaires, and facilitates understanding and correct coding of drug doses and past illnesses.

#### External validation

All events that result in hospitalization, and medical events deemed important, are validated by obtaining medical confirmation after obtaining consent from the participants. Evidence in support of putative events is classified according to quality, with the highest quality score being given to events directly supported by medical records, and lower quality scores to events reported by the patient's physician or by convincing interviews using the standardized protocols.

In the USA, patients may have many physicians, such that aspects of health status may not be known or recorded by most of them. For example, a patient being followed up by a rheumatologist for RA may be unaware of the details of gynaecological or chronic obstructive pulmonary disease issues that are treated by other physicians. Hospital records and details, even those related to allied issues such as joint surgery or infections, may also be unavailable to the primary rheumatologist. While it is

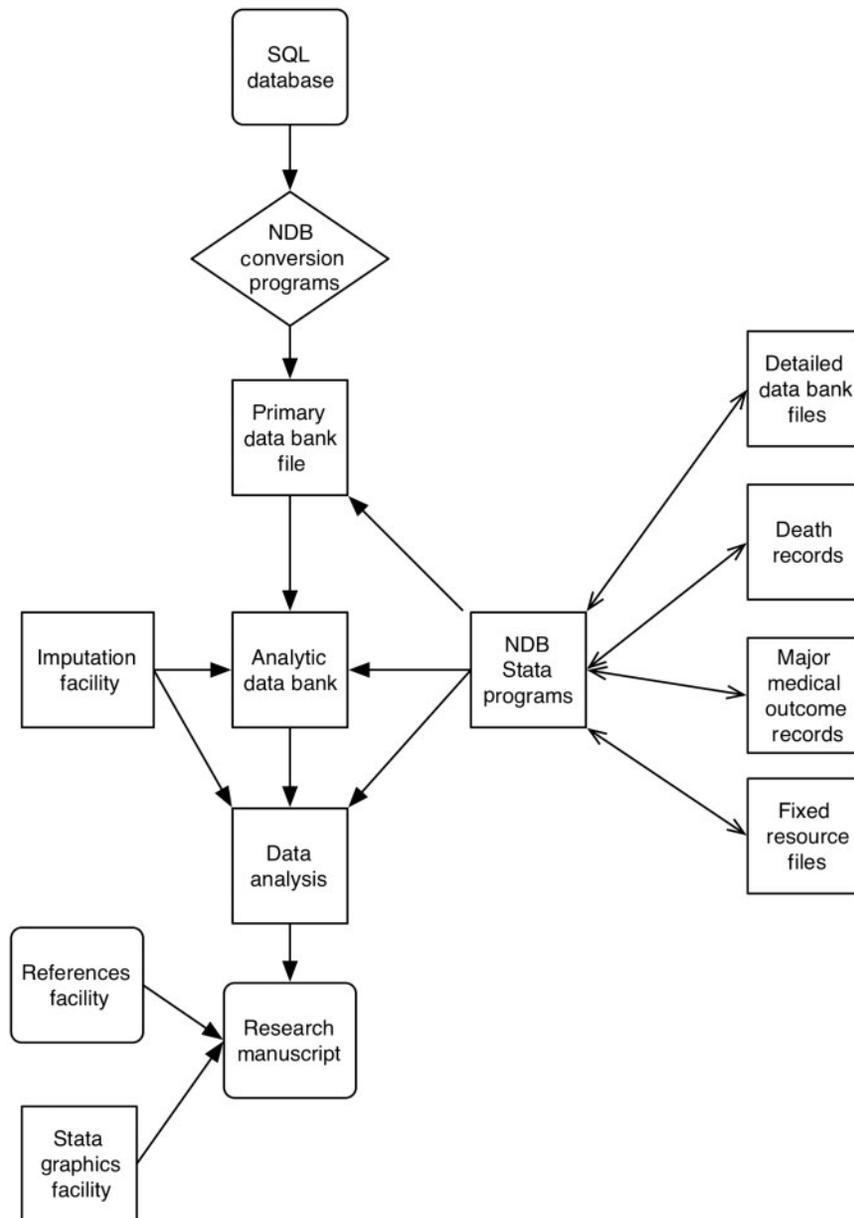
generally easy to obtain validation for inpatient events, it is much more difficult to obtain written results of outpatient consultations. In the NDB's experience, requests for outpatient records from orthopaedists, neurologists and dermatologists, for example, are often ignored. The NDB frequently uses short, faxed inquiries for rheumatologists to obtain key information. Asking simple, often checkbox-type questions by fax/return fax is an effective way to obtain cooperation from busy physicians and their staff.

#### Creation of research databases

The NDB maintains two databases (Fig. 3). The first is an SQL database. This database holds raw and converted data, and is used by the NDB staff to collect, organize and validate data. SQL databases, while excellent for maintenance of data, are difficult to use for analysis. Every night the NDB runs a series of complex programs that convert the SQL database into a research database suitable for statistical analysis (Fig. 3). This update brings each day's new work into the purview of the data analyst. The specially structured research database is a unique aspect of the NDB program.

The created research database uses the format and language of the Stata statistical software package [1]. Created by programs written by the authors, the programs reformat the data so that the analyst can use the data sets immediately. The programs score questionnaires and bring common variables into a single database. They organize time-related variables and structure the data set as needed for longitudinal analysis; and the research DB has a built-in missing data facility. As an example of how analysis might work, a researcher who is interested in

Fig. 3 NDB research DB structure and process.



mortality as an outcome can simply issue the 'getdeath' command and the programs will merge in all all-cause and cause-specific mortality data based on the US National Death Index (NDI) in a few seconds. The ease of obtaining such information occurs because all of the programming required is done behind the scenes during the research database creation process.

The NDB programs guarantee to the user that all data checking and organization have been done previously, and that the NDB stands behind the integrity of the data. The central difference between the NDB data structure and the usual data structure is that NDB data have been organized specifically for simple use by the analyst. An example of the use of NDB data can be found in the

video at <http://ndbresearch.blip.tv/>. In this link, we demonstrate a simple research analysis to examine the risk of glucocorticoids on the development of cataracts.

### The content of the multi-registry NDB

Although treatment may be disease specific, most symptoms and outcomes are disease independent. This allows a questionnaire to be useful across major rheumatic disorders. One advantage of the disease-independent questionnaire method is that it provides natural controls. For example, the rate of cardiovascular disease in RA can be compared with the rates in non-inflammatory disorders. In some instances questionnaires can differ by disease,

while still keeping 80% of common variables. For example, the NDB uses a shorter questionnaire in OA because specific questions regarding biologic therapy are not needed. The NDB lupus questionnaire contains lupus-specific items. In addition, it is possible to add disease-specific questionnaires into such shorter questionnaires, such as the OA WOMAC or the lupus systemic lupus activity questionnaire (SLAQ) or Lupus Damage Index Questionnaire.

Demographic characteristics, as shown in Table 1, are often key predictors and moderators of rheumatic disease outcomes, and should be an essential part of all assessments. Factors such as age, smoking, employment, work disability, marital status and household income change over time and repeat assessments are needed. Such data are usually more reliable when obtained by questionnaire than by perusal of physicians' records.

Questionnaires offer the ability to probe patient's health status in detail using longer questionnaires than are suitable for the clinic. In addition to the ubiquitous short form (SF)-36, the NDB collects the EQ-5D, the Widespread Pain Index (WPI), rheumatoid arthritis disease activity index (RADAI), the Medical Research Council (MRC) dyspnoea scale and various standard measures of function, among other scales (Table 1). By collecting somatic symptoms, it is possible to identify particular symptoms as well as obtain a measure of the degree of somatic symptom reporting.

The NDB follows up on all hospitalizations. We ask patients to provide dates and locations of hospitalization, and we contact patients for additional information when required. Although hospitalization data do identify most of the major events, there are a series of medical events that

may be determined by outpatient investigation and diagnosis. So we ask specific questions about events of interest, and we follow up as needed. For example, we inquire about all infection, malignancies and special events of interest such as the development of lupus or multiple sclerosis. The results of these determinations are placed in the major medical outcomes file (Fig. 3).

The determination of therapy by self-report presents challenges and rewards. We are able to capture non-prescription drugs and treatments for non-rheumatic conditions. In addition, this allows for capture of treatments that are prescribed by non-rheumatologists. Regardless of the treatment, we capture dose, frequency and start and stop dates. When treatments are discontinued, we capture the reason for discontinuation. If there are side effects, we capture the specific side effect, its severity and consequences.

In addition to capturing therapy, we determine the number and type of outpatient visits, and the utilization of imaging and laboratory tests. When combined with hospitalization data, the NDB is able to capture actual costs of therapy. A simple NDB program 'getcosts' brings such data immediately to the analyst. We keep a record of all deaths, and we search the US NDI annually for all patients no longer being followed.

The NDB has been used for the development and validation of a series of questionnaires. The DB can be used to test different versions of questionnaires. The NDB has been used in the development of the HAQ-II, Patient Activity Scale (PAS and PAS-II), Short Arthritis Scale (SAS), WPI, fibromyalgiansess scale, brief OA index and comorbidity index.

TABLE 1 The content of the NDB

Category	Items
Demographics and habits	Age sex, education, marital status, household income, specific employment, smoking, alcohol use and BMI
Comorbidity	Current and past comorbid conditions
Health status	
General	SF-36
Health utilities	EQ-5D EQ-VAS
Function	HAQ, HAQ-II, VAS function and SF-36 function
Quantitative symptom assessment	Pain, global, fatigue, sleep, cognitive function, health satisfaction, WPI, RADAI, RA activity, stiffness and MRC dyspnoea index
Somatic symptoms	Forty-seven specific symptoms
Special symptoms	Vasculitis
Work and disability	Employment and disability status, hours and days lost from work
Hospitalization	Dates and records of all hospitalizations
Specific serious medical conditions	Infections, malignancies, heart rhythm disturbances, renal failure, gastrointestinal ulcers, multiple sclerosis, fractures, etc.
Joint replacement surgery	Type, location and date
Immunizations	Influenza, pneumonia and herpes zoster
Anti-rheumatic therapy	Dose, duration (start/stop dates) for all therapies
Non-rheumatic therapy	Dose, duration (start/stop dates) for all therapies
Adverse effects of therapy	Description, severity and consequences
Utilization of services	Hospitalization, outpatient care and treatments, out of pocket expenses, insurance
Mortality	From alternative patient contact and NDI follow-up
Special items	Single or short-term questions

## Issues in the analysis of observational data

### DB bias

DBs that depend on volunteer patients or volunteer physicians are biased about characteristics of patients in general. Patients in the USA seeing rheumatologists are less likely to be members of ethnic minorities or to have their income below the median of all rheumatic disease patients. In addition, participants in surveys usually have better education and other social class advantages compared with the general population. These differences may create problems when DB outcomes are compared with those in the general population. However, covariate control may sometimes help to overcome that problem when adjustments are made for age, sex, education, income, ethnicity, smoking and related items.

### Missing data

Observational DBs also suffer from missing data in that patients may omit or not complete some questionnaires. NDB addresses this issue by follow-up with patients for key missing variables. For example, missing data regarding infections, hospitalization and malignancies always result in additional contact with the patient. For most clinical variables, missing data are not much of a problem. For example, only 1.1% of HAQ scores and 6.4% of SF-36 PCS scores were missing in NDB comprehensive questionnaires. Where necessary, the NDB uses multiple imputation by chained equations to impute missing variables [2, 3], and this methodology is built in to the NDB research data set. Attrition is a problem in all observational studies [4]. The NDB conducts exit interviews with patients who withdraw consent, and incorporates such data in analyses.

### Misclassification

Misclassification of events can be a problem when events reported by patients or elicited by physicians cannot be validated. For example, when the patient tells the rheumatologist or reports on an NDB questionnaire that he has had pneumonia, then that event must be validated. Hospital records (Class I evidence) may not be available. In that instance, physician contact may be acceptable (Class II evidence). In addition, the patient may report convincing details of the hospitalization to a trained interviewer who uses a standardized assessment protocol (Class II evidence). The NDB classifies evidence by quality, and does not use evidence below Class II. Under-identification of events, as in the case of true events that do not reach the Class II level, can result in underestimation of incidence rates. The use of internal controls (e.g. RA vs non-inflammatory disorders) can sometimes aid in the proper reporting of rate ratios.

The NDB also deals with 'soft' events. By soft events we mean those events that occur in the outpatient setting (usually not in rheumatology practice) for which interpretation of the events places a role in their classification.

Examples of such events include reports of herpes zoster, retinal toxicity from HCQ, development of lupus or multiple sclerosis. Such events are usually substantially over-reported by patients. To validate events such as these, two-stage procedures are necessary. For herpes zoster, we re-contacted patients and interviewed them using validated research questionnaires [5]. For HCQ toxicity, we sought copies of medical records from eye specialists and had them interpreted by experts [6].

### Bias and analysis

A series of biases can afflict longitudinal observational DB studies, most often when the purpose of the study is to determine the effect of treatment. Usually, the most important problem is confounding: when factors that lead to the selection of treatment also influence the outcome of treatment. Confounding occurs because of non-random selection of treatment, leading to the state where patients who receive the treatment are systematically different from those who do not receive the treatment.

Although such biases are sometimes theoretical and difficult to document, they can also be quite dramatic. As part of longitudinal DB activities, the NDB studied 6637 RA and OA patients in 1998 just before and just after the introduction of COX-2 inhibitors [7]. Patients starting a new COX-2-specific inhibitor had a greater lifetime history of adverse reactions, more severe scores for pain, functional disability, fatigue, helplessness and global severity, and used more inpatient and outpatient services than patients who would not switch to COX-2-specific inhibitors. In addition, they had worse outcomes after switching.

The bias and the variables related to it described above had certain characteristics: observability and stability with respect to time. That is, the direction and extent of the bias did not change with relative or calendar time. In this example, the effect was stable because the treatment effect was measured only once. There is a relatively straightforward statistical solution to the bias just described, adjusting for differences in baseline covariates. This is usually done with a propensity score, a method of obtaining a single score that describes the baseline differences between patients. Propensity scores should be determined based on variables that influence selection of the treatment and are associated with the outcome of the treatment.

Propensity scores are effective when, and only when, they account for all key covariates that influence treatment selection. But there are often unobserved covariates that play a role. Patients may receive treatments because their general health is better, they are better able to pay for treatments (USA), are of higher or lower social class, a pharmaceutical company compensates physicians for prescription (USA), or they have different medical insurance coverage, among many other confounders. Thus, it is possible to match patients with propensity scores on observed covariates yet fail to match them on unobserved covariates, with the result that the outcome of the study is biased despite statistical manipulation; and it is often

TABLE 2 Categories of NDB DB research<sup>a</sup>

Category	General	RA	FMS	Lupus	OA
Programming and statistical analyses		[8]			
Questionnaire development and validation	[9, 10]	[10–15]	[16–18]	[18–20]	[21, 22]
Criteria development and testing		[23, 24]	[25]		
Adverse effects and comorbid outcomes	[26]	[6, 27–32]	[33]	[6]	
The course of illness		[33–38]	[38, 39]		[35]
Disease state descriptions and interactions	[40]	[33, 41–43]	[33, 43]		[43, 44]
Costs, cost-effectiveness and work disability		[13, 45–47]			
Treatment		[32, 42, 48–50]			

<sup>a</sup>Selected references since 1998 from 216 published articles.

impossible to gauge the direction and extent of the bias. Propensity scores, while often used, should not give one a sense of security that control is adequate.

Confounding also occurs over time. At the simplest level, indication for the use of drugs may change over time. Biologics may be prescribed to patients with the worst outcomes earlier in the life course of the treatment and later to those with much better outcomes. Similarly, in the USA, insurance coverage and availability of treatment can change with time. Consequently, a HAQ or pain score may be seen as ‘high’ at one time and ‘low’ at another time in the life of the longitudinal DB. These issues make it difficult to apply propensity scores and to adequately ‘balance’ them in longitudinal DBs. Still a third type of problem occurs when the changes in the observed outcome over time influences treatment use.

There are several approaches to DB analysis in settings such as described above. First, it may be wise to avoid analyses where there cannot be adequate covariate control. In addition, certain analyses are not substantially influenced by treatment. DBs can do many things; it is not necessary for them to do all things. Second, it should be remembered that developing adequate models requires substantial exploratory analysis. The analyst should develop many models and observe how they differ. Probing for bias and estimating its extent and the likely effect is important. We would also urge caution when accepting the results of observational studies where pharmaceutical companies have a strong role in the design and analysis of studies. It is always possible to find a positive result if the question is posed the right way or if the sample is special.

### The ideal DB

The ideal DB contains patient, physician, laboratory and imaging data, and is linked to national disease and death registries. It should have an unbiased selection of patients. The closest we currently come to such DBs are the Swedish registries. Problems that often preclude such a model include inadequate funding and privacy laws. The UK experiment of mandating a biologics registry as a condition of biologic approval offers a methodology for outcome studies. The UK method is particularly praiseworthy, because its analyses are independent of the

pharmaceutical industry, although the funding for the projects comes from industry.

We would argue, however, that single outcome, limited purpose registries are insufficient to address the issues of rheumatic disease outcomes. Biologics registries, for example, address only limited issues in a minority of RA patients.

There are cost savings in using large observational DBs compared with registries that require large physician and staff time or the use of administrative data that lack patient-level detail. While the NDB model is less than ideal, for the costs it helps to answer important questions and provides a substantial resource for the world medical community (patients, physicians and society).

### NDB projects and the uses of DBs and registries

DBs have the ability to address multiple important issues in rheumatic diseases (Table 2). While non-random assignment to treatment causes difficulty in discerning treatment effect, the naturalistic data of DBs and registries provide just the right setting for non-treatment effect studies, particularly when the characteristics of patients in the DB approximate those in the community. The NDB has used its DB to aid in the development and validation of a series of questionnaires across the spectrum of rheumatic diseases. The ability to address multiple diseases is important because many questionnaires are not disease specific, and there is a natural desire to understand how questionnaires may work across illnesses. For example, the HAQ, HAQ-II, SF-36, EQ-5D and fibromyalgians scales are not disease specific. DBs are the perfect environment for questionnaire development because so many covariates are also available. We have also used DB patients and facilities to explore the development of FM criteria.

DBs are also an appropriate setting for assessing the risk of development of comorbid conditions, for example, cardiovascular and malignant disease. They are also very useful for determining the rates and predictors of direct and indirect costs, work disability and mortality.

In summary, the NDB provides a cost-effective, cross-disease method to explore critical issues for patients, physicians and society. The NDB is simple to use and

provides useful tools for research, teaching and data analysis.

#### Rheumatology key messages

- The NDB for rheumatic diseases is a patient-based multi-disease, multi-purpose longitudinal rheumatic disease registry.
- Rheumatologist-written programs make NDB data immediately available to analysts.

*Disclosure statement:* F.W. is an employee of the NDB. The NDB has received research grants from Bristol Myers Squibb, UCB and Pfizer. The other author has declared no conflicts of interest.

#### References

- 1 Stata Corporation. Stata Statistical Software: Release 11.0. College Station, TX: Stata Corporation, 2009.
- 2 Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol* 2010;63:205–14.
- 3 Royston P. Multiple imputation of missing values. *Stata J* 2004;4:227–41.
- 4 Twisk J, de Vente W. Attrition in longitudinal studies. How to deal with missing data. *J Clin Epidemiol* 2002;55:329–37.
- 5 Wolfe F, Michaud K, Chakravarty EF. Rates and predictors of herpes zoster in patients with rheumatoid arthritis and non-inflammatory musculoskeletal disorders. *Rheumatology* 2006;45:1370–5.
- 6 Wolfe F, Marmor MF. Rates and predictors of hydroxy-chloroquine retinal toxicity in patients with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Care Res* 2010 (in press).
- 7 Wolfe F, Flowers N, Burke TA, Arguelles LM, Pettitt D. Increase in lifetime adverse drug reactions, service utilization, and disease severity among patients who will start COX-2 specific inhibitors: quantitative assessment of channeling bias and confounding by indication in 6689 patients with rheumatoid arthritis and osteoarthritis. *J Rheumatol* 2002;29:1015–22.
- 8 Fewell Z, Hernan MA, Wolfe F, Tilling H, Choi HK, Sterne AC. Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004;4:402–20.
- 9 Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. *J Rheumatol* 2004;31:1896–902.
- 10 Wolfe F, Michaud K, Pincus T. Development and validation of the health assessment questionnaire II: a revised version of the health assessment questionnaire. *Arthritis Rheum* 2004;50:3296–305.
- 11 Wolfe F, van der Heijde DM, Larsen A. Assessing radiographic status of rheumatoid arthritis: introduction of a short erosion scale. *J Rheumatol* 2000;27:2090–9.
- 12 Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol* 2001;28:982–9.
- 13 Walker N, Michaud K, Wolfe F. Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. *J Rheumatol* 2005;32:1006–12.
- 14 Wolfe F, Michaud K, Pincus T. Preliminary evaluation of a visual analog function scale for use in rheumatoid arthritis. *J Rheumatol* 2005;32:1261–6.
- 15 Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies and clinical trials: the patient activity scale (PAS/PAS-II). *J Rheumatol* 2005;32:2410–5.
- 16 Wolfe F, Hawley DJ, Goldenberg DL, Russell IJ, Buskila D, Neumann L. The assessment of functional impairment in fibromyalgia (FM): Rasch analyses of 5 functional scales and the development of the FM health assessment questionnaire. *J Rheumatol* 2000;27:1989–99.
- 17 Wolfe F. Pain extent and diagnosis: development and validation of the regional pain scale in 12,799 patients with rheumatic disease. *J Rheumatol* 2003;30:369–78.
- 18 Wolfe F, Petri M, Alarcon GS *et al*. Fibromyalgia, systemic lupus erythematosus (SLE), and evaluation of SLE activity. *J Rheumatol* 2009;36:82–8.
- 19 Pons-Estel B, Sánchez-Guerrero J, Romero-Díaz J *et al*. Validation of the Spanish, Portuguese and French versions of the lupus damage index questionnaire: data from North and South America, Spain and Portugal. *Lupus* 2009;18:1033–52.
- 20 Costenbader KH, Khamashta M, Ruiz-Garcia S *et al*. Development and initial validation of the self-assessed Lupus Damage Index Questionnaire (LDIQ). *Arthritis Care Res* 2010;62:559–68.
- 21 Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Ann Rheum Dis* 1999;58:563–8.
- 22 Wolfe F, Michaud K, Kahler K, Omar M. The short arthritis assessment scale: a brief assessment questionnaire for rapid evaluation of arthritis severity in research and clinical practice. *J Rheumatol* 2004;31:2472–9.
- 23 Wolfe F, Rasker JJ, Boers M, Wells GA, Michaud K. Minimal disease activity, remission, and the long-term outcomes of rheumatoid arthritis. *Arthritis Rheum* 2007;57:935–42.
- 24 Shaver TS, Anderson JD, Weidensaul DN *et al*. The problem of rheumatoid arthritis disease activity and remission in clinical practice. *J Rheumatol* 2008;35:1015–22.
- 25 Wolfe F, Clauw D, Fitzcharles MA. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care Res* 2010;62:600–10.
- 26 Wolfe F, Anderson J, Burke TA, Arguelles LM, Pettitt D. Gastroprotective therapy and risk of gastrointestinal ulcers: risk reduction by COX-2 therapy. *J Rheumatol* 2002;29:467–73.
- 27 Wolfe F, Michaud K. The loss of health status in rheumatoid arthritis and the effect of biological therapy: a longitudinal observational study. *Arthritis Res Ther* 2010;12:835–47.
- 28 Wolfe F, Michaud K. Lymphoma in rheumatoid arthritis: the effect of methotrexate and anti-tumor necrosis factor

- therapy in 18,572 patients. *Arthritis Rheum* 2004;50:1740–51.
- 29 Wolfe F, Michaud K, Anderson J, Urbansky K. Tuberculosis infection in patients with rheumatoid arthritis and the effect of infliximab therapy. *Arthritis Rheum* 2004;50:372–9.
  - 30 Wolfe F, Michaud K. The risk of myocardial infarction and pharmacologic and nonpharmacologic myocardial infarction predictors in rheumatoid arthritis: a cohort and nested case-control analysis. *Arthritis Rheum* 2008;58:2612–21.
  - 31 Wolfe F, Michaud K. Biologic treatment of rheumatoid arthritis and the risk of malignancy: analyses from a large US observational study. *Arthritis Rheum* 2007;56:2886–95.
  - 32 Wolfe F, Michaud K. The effect of methotrexate and anti-tumor necrosis factor therapy on the risk of lymphoma in rheumatoid arthritis in 19,562 patients during 89,710 person-years of observation. *Arthritis Rheum* 2007;56:1433–9.
  - 33 Wolfe F, Michaud K. Prevalence, risk, and risk factors for oral and ocular dryness with particular emphasis on rheumatoid arthritis. *J Rheumatol* 2008;35:1023–30.
  - 34 Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. *Arthritis Rheum* 2000;43:2751–61.
  - 35 Wolfe F, Hawley DJ. The comparative risk and predictors of adverse gastrointestinal events in rheumatoid arthritis and osteoarthritis: a prospective 13 year study of 2131 patients. *J Rheumatol* 2000;27:1668–73.
  - 36 Choi HK, Hernan MA, Seeger JD, Robins JM, Wolfe F. Methotrexate therapy and mortality in patients with rheumatoid arthritis: a prospective study. *Lancet* 2002;359:1173–7.
  - 37 Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. *Arthritis Rheum* 2003;48:1530–42.
  - 38 Wolfe F, Michaud K. Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize RA patients with fibromyalgia. *J Rheumatol* 2004;31:695–700.
  - 39 Wolfe F, Michaud K. Outcome and predictor relationships in fibromyalgia and rheumatoid arthritis: evidence concerning the continuum versus discrete disorder hypothesis. *J Rheumatol* 2009;36:831–6.
  - 40 Hawley DJ, Wolfe F, Lue FA, Moldofsky H. Seasonal symptom severity in patients with rheumatic diseases: a study of 1,424 patients. *J Rheumatol* 2001;28:1900–9.
  - 41 Wells GA, Boers M, Shea B *et al*. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;32:2016–24.
  - 42 Wolfe F, Michaud K. The loss of health status in rheumatoid arthritis and the effect of biologic therapy: a longitudinal observational study. *Arthritis Res Ther* 2010;12:R35.
  - 43 Bergman MJ, Shahouri SS, Shaver TS *et al*. Is fatigue an inflammatory variable in rheumatoid arthritis (RA)? Analyses of fatigue, osteoarthritis, and fibromyalgia. *J Rheumatol* 2009;36:2788–94.
  - 44 Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. *Rheumatology* 1999;38:355–61.
  - 45 Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in patients with rheumatoid arthritis: a three-year study of 7,527 patients. *Arthritis Rheum* 2003;48:2750–62.
  - 46 Allaire S, Wolfe F, Niu J, LaValley M, Michaud K. Work disability and its economic effect on 55-64-year-old adults with rheumatoid arthritis. *Arthritis Rheum* 2005;53:603–8.
  - 47 Wolfe F, Michaud K, Choi HK, Williams R. Household income and earnings losses among 6,396 persons with rheumatoid arthritis. *J Rheumatol* 2005;32:1875–83.
  - 48 Dewitt EM, Glick HA, Albert DA, Joffe MM, Wolfe F. Medicare coverage of tumor necrosis factor alpha inhibitors as an influence on physicians' prescribing behavior. *Arch Intern Med* 2006;166:57–63.
  - 49 Wolfe F, Michaud K. Assessment of pain in rheumatoid arthritis: minimal clinically significant difference, predictors, and the effect of anti-tumor necrosis factor therapy. *J Rheumatol* 2007;34:1674–83.
  - 50 Wolfe F, Michaud K. Resistance of rheumatoid arthritis patients to changing therapy: discordance between disease activity and patients' treatment choices. *Arthritis Rheum* 2007;56:2135–42.