



Reyher, K. K., Dufour, S., Barkema, H. W., Des Côteaux, L., Devries, T. J., Dohoo, I. R., Keefe, G. P., Roy, J-P., & Scholl, D. T. (2011). The national cohort of dairy farms - a data collection platform for mastitis research in Canada. *Journal of Dairy Science*, 94(3), 1616-26.  
<https://doi.org/10.3168/jds.2010-3180>

Publisher's PDF, also known as Version of record

Link to published version (if available):  
[10.3168/jds.2010-3180](https://doi.org/10.3168/jds.2010-3180)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



## The National Cohort of Dairy Farms—A data collection platform for mastitis research in Canada

K. K. Reyher,\*†<sup>1</sup> S. Dufour,\*‡ H. W. Barkema,\*§ L. Des Côteaux,\*# T. J. DeVries,\*|| I. R. Dohoo,\*† G. P. Keefe,\*† J.-P. Roy,\*# and D. T. Scholl\*‡

\*Canadian Bovine Mastitis Research Network, C.P. 5000, St-Hyacinthe, Québec, Canada J2S 7C6

†Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada C1A 4P3

‡Département de Pathologie et Microbiologie, Faculté de Médecine Vétérinaire, Université de Montréal, C.P. 5000, St-Hyacinthe, Québec, Canada J2S 7C6

§Department of Production Animal Health, Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada T2N 4N1

#Département de Sciences Cliniques, Faculté de Médecine Vétérinaire, Université de Montréal, C.P. 5000, St-Hyacinthe, Québec, Canada J2S 7C6

||Department of Animal and Poultry Science, University of Guelph, Kemptville Campus, Kemptville, Ontario, Canada K0G 1J0

### ABSTRACT

Costs and feasibility of extensive sample collection and processing are major obstacles to mastitis epidemiology research. Studies are often consequentially limited, and fundamental mastitis researchers rarely have the opportunity to conduct their work in epidemiologically valid populations. To mitigate these limitations, the Canadian Bovine Mastitis Research Network has optimized research funds by creating a data collection platform to provide epidemiologically meaningful data for several simultaneous research endeavors. This platform consists of a National Cohort of Dairy Farms (NCDF), Mastitis Laboratory Network, and Mastitis Pathogen Culture Collection. This paper describes the implementation and operation of the NCDF, explains its sampling protocols and data collection, and documents characteristics, strengths and limitations of these data for current and potential users. The NCDF comprises 91 commercial dairy farms in 6 provinces sampled over a 2-yr period. Primarily Holstein-Friesian herds participating in Dairy Herd Improvement milk recording were selected in order to achieve a uniform distribution among 3 strata of bulk tank somatic cell counts and to reflect regional proportions of freestall housing systems. Standardized protocols were implemented for repeated milk samplings on clinical mastitis cases, fresh and randomly selected lactating cows, and cows at dry-off and after calving. Just fewer than 133,000 milk samples were collected. Demographic and production data were recorded at individual cow and farm levels. Health management data are documented and extensive questionnaire data detailing farm management and cleanliness information are also captured.

The Laboratory Network represents coordinated regional mastitis bacteriology laboratories using standardized procedures. The Culture Collection archives isolates recovered from intramammary infections of cows in the NCDF and contains over 16,500 isolates, all epidemiologically cross-referenced between linked databases. The NCDF is similar to Canadian dairies in relation to mean herd size, average production, and freestall percentages. Pathogen recovery was greater than anticipated, particularly for coagulase-negative staphylococci and *Corynebacterium* spp. International scientists are encouraged to use this extensive archive of data and material to enhance their own mastitis research.

**Key words:** udder health, mastitis, cohort, mastitis isolate

### INTRODUCTION

Although considerable effort has been expended in attempting to delineate its causes and control over the past several decades, mastitis is still the most common production disease on dairies worldwide, and the most costly (Seegers et al., 2003; Halasa et al., 2007). It is a multi-faceted disease complex of the global dairy industry and concern for its prevention and treatment still exists.

Numerous studies yielding important progress have been conducted and publicized with regard to understanding and controlling mastitis. Many of these studies, however, are regionally focused and relatively few are longitudinal. Longitudinality permits estimation of the incidence of new IMI and elimination rates, as well as the in-depth study of risk factors, whereas cross-sectional designs restrict attention to prevalent infection (Rothman et al., 2008). Research on a large scale is often prohibitively expensive and logistically challenging, requiring multiple visits to widely distributed farms and massive numbers of repeated milk samplings

Received February 17, 2010.

Accepted November 30, 2010.

<sup>1</sup>Corresponding author: kreyher@upei.ca

involving considerable manpower and laboratory work. Most recent large longitudinal studies focus on clinical mastitis (Barkema et al., 1998; Bradley et al., 2007; Olde Riekerink et al., 2008), whereas larger studies detailing IMI over the dry period were conducted several decades ago (Funk et al., 1982; Browning et al., 1994). Studies on IMI in apparently normal lactating cows are often cross-sectional in design, and rare longitudinal studies are most often performed on limited populations (Zadoks et al., 2001). Although these studies add to our knowledge of the incidence and prevalence of IMI on dairy farms, the lack of large-scale, longitudinal studies may not provide a fully informed appreciation of the challenges associated with the improvement of udder health.

The Canadian Bovine Mastitis Research Network (CBMRN) has attempted to decrease data collection costs for individual research projects as well as enable researchers to study a national-level target population by establishing a single national-level longitudinal cohort for mastitis research. Coordination of a common data collection effort for multiple research lines was developed to ensure uniformity and rigor of data quality. The ability to efficiently share data among collaborators from different research domains was also a focus, and the suitability of collected data for all intended research purposes was given considerable thought before the creation of a shared data collection platform.

The objectives of this paper are 3-fold: to describe the implementation and operation of the National Cohort of Dairy Farms (NCDF), to explain the sampling protocols and data collection occurring within the NCDF, and to document characteristics, strengths, and limitations of these data for current and potential users.

## MATERIALS AND METHODS

### Overview

The core research platform of the CBMRN consists of 3 parts: the NCDF, the Mastitis Laboratory Network (MLN), and the Mastitis Pathogen Culture Collection (MPCC). The NCDF is a sample of herds from 6 Canadian provinces followed over 2 yr, which serves to decrease the incremental costs of epidemiological studies and overcome regional discontinuities by providing national-level data on commercial dairy farms. The MLN provides coordinated milk bacteriology analyses across the national cohort. The MPCC archives geographically representative and epidemiologically referenced IMI isolates as a resource for researchers requiring isolates with particular features or from specified subpopulations of dairy farms or pathogens with particular features.

### Organizational Structure

Research centers in Alberta, Ontario, Quebec, and the Atlantic provinces (Prince Edward Island, Nova Scotia, and New Brunswick) were established by regional directors of the CBMRN. Farms in these 6 provinces were then identified and enrolled if they agreed to cooperate with and participate directly in data collection. Because Canada has 2 official languages, documents and questionnaires meant for producers were translated from French into English, or vice versa (Dufour et al., 2010).

### Herd Selection

The NCDF is intended to represent as closely as is feasible the Canadian commercial dairy farm population. An overall target of 80 farms was expected to be realized by an assumed attrition of 20% of 100 initially enrolled farms.

Farm selection took place in 2 stages. Farms from the 4 regions were enrolled with the following targets: 20 farms in Alberta, 30 farms in Ontario, 30 farms in Quebec, and 20 farms in the Atlantic region. Within the regions, farms that were willing to participate were recruited to achieve specified distributions of bulk tank SCC (BTSCC) and housing systems. Based on 2006 average BTSCC, herds were selected to achieve a regionally uniform distribution of 3 BTSCC ranges ( $\leq 150,000$  cells/mL, 150,000 to 300,000 cells/mL, and  $> 300,000$  cells/mL). The target distribution of housing systems was obtained from a recent mail survey of Canadian dairy farm characteristics (Olde Riekerink et al., 2010; Table 1), and herds were enrolled to match a proportion of freestall systems within 15 percentage points of their respective provincial freestall percentages. Furthermore, the farms had to be close enough to the coordinating center to make weekly farm visits feasible. All farms had  $> 80\%$  Holstein-Friesian cows, were milking twice per day at the time of enrollment, and subscribed to a regular regional DHI recording system.

### Milk Sample Collection

Standardized sample taking, sample handling and data collection protocols were published in a reference manual (available upon request), and technicians trained producers to follow the set protocols. Three individual-cow milk-sampling series were performed that focused on 1) clinical mastitis, 2) nonclinical lactating cows, and 3) cows before dry-off and after calving. Monthly bulk tank milk samples were also taken. At least 30 mL of milk was taken for each milk sample to ensure adequate volume for bacteriological culture and somatic cell testing.

**Table 1.** Actual 2006 12-mo average bulk tank SCC (cells/mL), freestall percentages, and participation results of farms enrolled in the National Cohort of Dairy Farms (NCDF)

Characteristic	Region			
	Alberta	Ontario	Québec	Atlantic Canada
Bulk tank SCC <150,000	5	10	3	3
Bulk tank SCC 150,000–300,000	9	10	18	11
Bulk tank SCC >300,000	3	9	6	4
Target proportion of freestall farms <sup>1</sup>	0.79	0.15	0.30	0.47
Achieved proportion of freestall farms	0.81	0.09	0.28	0.48
2007				
Enrolled	17	27	29	18
Completed	17	27	28	18
2008				
Started	15	23	25	16
Completed	12	23	25	16

<sup>1</sup>Olde Riekerink et al. (2010).

For clinical cases of mastitis, farmers were asked to sample cows they identified as having abnormal milk or more severe clinical signs. Single samples of only the affected quarter were taken on the day of diagnosis, and repeated at 2 to 3 wk, and again at 4 to 5 wk after diagnosis. A mastitis clinical score was given to each sample, categorizing the quarter as having abnormal milk only (mastitis score = 1); abnormal milk plus a swollen quarter (mastitis score = 2); or abnormal milk, swollen quarter, and sick cow (mastitis score = 3; Sears and McCarthy, 2003). The cow's rectal temperature at the time of diagnosis, the date the milk from the affected quarter first appeared normal, and the date the producer resumed putting milk from the affected quarter into the bulk tank were also recorded.

In the first year, farmers were asked to manage cases of clinical mastitis as was normal practice on their farm. In the second year, a subset of farms (31) participated in an on-farm culture and treatment project and adopted treatment as specified by the project. Monetary incentives (CAN\$30 per completed record and follow-up samples for each clinical mastitis case) were provided to farmers enrolled in this project.

The second sampling series was longitudinal incidence-density sampling directed at lactating cows that were considered clinically normal (no abnormal milk) at the beginning of the series. Independent sampling periods, consisting of 3 visits separated by intervals of 3 wk, were established during the spring of 2007 (March–May), winter of 2008 (January–March), and summer of 2008 (June–August). An intensive sampling period also occurred during the summer of 2007 (June–August), which consisted of 7 weekly samplings. Quarter samples were collected by technicians from 15 cows per period: 10 chosen randomly from a sampling frame of cows expected to remain in the herd for at least 2 mo from

the initialization of sampling, along with the 5 cows that had most recently calved. This sample size was predetermined to permit estimation of incidence rates of IMI on NCDF farms to within 0.4 cases per 100 cows per month. Additionally, composite samples were taken for future analyses at the first sampling of each period. Teat end scores were recorded at the first and last sampling of each period using the scoring system of Neijenhuis et al. (2000), so that each cow had a repeat teat end score. Any quarter that had mastitis during the sampling scheme was noted and samples were collected for the clinical mastitis cases series.

The third sampling series was cows sampled before dry-off and again after calving. The first 15 cows to be dried off that were expected to remain in the herd until at least 2 wk after calving were enrolled in 2007. If a cow dropped out before the entire sequence of samples was taken, another cow was enrolled to replace the original cow, until a total of 15 cows with complete sampling series was achieved for each herd. Composite and individual quarter samples were collected (largely by farm personnel) between 4 and 2 wk before dry-off, between 2 wk and dry-off, within 24 h of calving, and from 1 to 2 wk after calving. Mastitis clinical scores were recorded for each quarter sampled.

### **Milk Sample Processing**

Milk samples taken by farm personnel were frozen on the farm until monthly pick-up by technicians. Samples taken by technicians were placed on ice and subsequently frozen for storage at  $-20^{\circ}\text{C}$ . Samples were submitted to 1 of 4 milk bacteriology laboratories comprising the MLN. Samples submitted to laboratories that were off-site were shipped frozen and on ice by overnight courier. To offer a rapid turnaround to pro-

ducers, clinical mastitis samples were cultured as soon as they arrived in the laboratories.

A standardized protocol for bacteriological culture and species identification was developed based on National Mastitis Council (NMC) guidelines (Hogan et al., 1999), and was strictly adhered to by participating laboratories. After completion of bacteriological culturing, samples were refrigerated or refrozen in preparation for transport for SCC analysis.

Isolates of growth that were considered significant were conserved in the MPCC (for details, please see National Cohort of Dairy Farms, 2009). A milk sample from which 3 or more different species were cultured was considered contaminated, although *Staphylococcus aureus* and *Streptococcus agalactiae* identified in these samples were enumerated and conserved. Colonies selected for conservation were resuspended in trypticase soy broth (TSB) and 15% glycerol in a 2-mL cryovial (Nalgene, Rochester, NY) labeled with the associated sample barcode. Inoculated cryovials were frozen and shipped overnight on ice to the CBMRN MPCC at the Université de Montréal. The conserved isolates are maintained at  $-80^{\circ}\text{C}$ .

All samples other than the initial clinical mastitis samples were analyzed for SCC. Samples were preserved with bronopol and shipped from bacteriology laboratories overnight on ice to Maritime Quality Milk at the University of Prince Edward Island. The SCC analysis was conducted using the Fossomatic method (Fossomatic 4000 series, Foss Electric, Hillerød, Denmark). When samples were not processed for SCC on the day they were received, they were either stored frozen (if they arrived frozen) or refrigerated and processed within 5 d (if they arrived thawed).

### Other Data Collection

Health data on all cattle from all farms were collected, with particular interest in incidents relating to udder health and antibiotic administration for any reason. Farmers differed in the methods by which they recorded health data; some recorded in a binder provided by the NCDF, whereas others left the technicians to transcribe or download information from another set of records [e.g., those used for Canadian Quality Milk (<http://www.dairygoodness.ca/en/media/organization/programs/canadian-quality-milk/canadian-quality-milk.htm>) or other pre-existing herd records]. Farmers recorded data continuously and were motivated by regular visits by technicians.

Three separate pre-tested questionnaires were administered by personal interview by regional technicians to all remaining producers during regular herd visits (Dufour et al., 2010). The first questionnaire was intended

to capture data on general housing, maternity pens, biosecurity, diseases, nutrition, and producer attitudes in place during the study period. The second questionnaire covered milking facilities, milking procedures, milking equipment, and management of clinical and subclinical mastitis used during the study period. Finally, the third questionnaire was solely dedicated to dry period management, addressing target dry period length, dry cow treatments, body condition scoring, and nutrition of dry cows, as well as dry cow housing. Specific questions on this questionnaire addressed each of the sampling years separately, whereas others addressed practices at the time of survey. Variables that could be observed on-farm were recorded directly by the regional technicians, who were also asked to provide subjective observations on general farm management. Observations related to milking procedures were recorded during lactational samplings (Dufour et al., 2010).

### Affiliated Studies

In addition to the standardized data and material collection performed for multi-distribution, supplementary data were collected for specific projects that are available to all researchers through collaborative arrangements. An innate immune responsiveness study was conducted, and a DNA archive was established on a subset of cows enrolled in the incidence density samplings. For a project designed to characterize antimicrobial drug use and mastitis pathogen antimicrobial susceptibility patterns on participating farms, and to validate antimicrobial treatments and general health records, producers were asked to dispose of all empty containers for drugs used on-farm (particularly antibiotic products) in separate bins. A project investigating the gastrointestinal parasite load of pastured cattle and response to eprinomectin was carried out on a subset of 40 NCDF farms, and a project investigating the relationship between feeding strategy, lying behavior, and IMI incidence was conducted on 6 herds. Species and strain distribution of acute and persistent CNS IMI is also being described.

### Data Entry, Storage, and Access

Data with a high risk of containing errors, such as records of the wrong cow being sampled, caused at the farm level, were entered regionally; all other data were entered at the central administrative center. All data (including DHI downloads and the MPCC holdings) are linked and stored on a central server in an Access relational database (Access 2003, Microsoft Corp., Redmond, WA). These databases are available to participating researchers by way of a password-protected web-



site ([www.mastitisnetwork.org](http://www.mastitisnetwork.org)). The reference manual of all standard operating procedures is also available at this site or can be made available by contacting the author.

Quality control checks were built into the data entry process to avoid errors. These check devices prevented dual entry of samples, monitored sample identifications for accuracy, and employed multiple-center entry for checking. Validation files were also created to validate a sampling series, and systematic routines identified suspicious data values and entry errors.

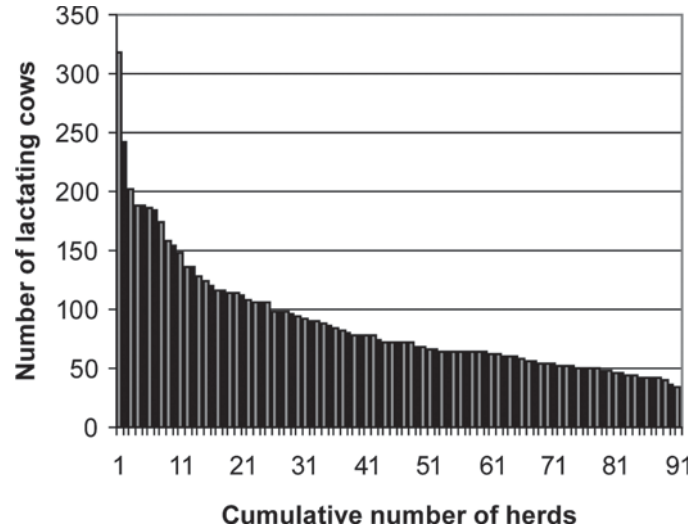
### Statistical Analyses

Mean herd size and SCC means were estimated using the average herd size of each farm at regular DHI tests. Milk production was calculated using the DHI-provided 305-d production for each herd. Unless otherwise stated, compliance was calculated according to the proportion of complete sets of samples taken for a prescribed sampling scheme. Both the geometric and production-weighted arithmetic mean SCC were calculated over the whole period of the 2-yr CBMRN data collection, using DHI data. Geometric means were calculated by taking the log of the SCC for each cow at each test, averaging these log-transformed SCC values, and then exponentiating them to come up with a mean for each herd for each test over the 2-yr period. Production-weighted arithmetic means were calculated using milk weights and SCC values from each cow at every test date. A weighting based on milk production for each cow at each test was calculated and used to calculate the contribution that cow would have made to the herd SCC with all her milk included in the bulk tank. This weight was then applied to the SCC measurement for each cow on each test date, and an average for each farm was calculated. All analyses were carried out using Stata/MP 11.0 (StataCorp, College Station, TX), with full details available upon request from the senior author.

## RESULTS

### Descriptive Characteristics

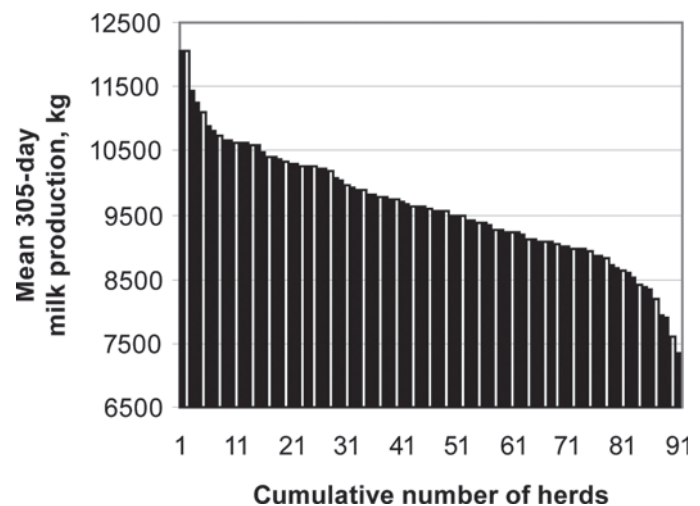
In late 2006 and early 2007, 91 dairy farms were recruited: 17 farms in Alberta, 27 in Ontario, 29 in Quebec, 5 in New Brunswick, 5 in Nova Scotia and 8 in Prince Edward Island (Table 1). Of the 91 farms enrolled, 90 completed the first year of data collection (2007). The farm that withdrew was unable to dedicate the effort necessary to complete sampling. For 2008, 11 of the least compliant farms were given the opportunity to cease sampling. Three additional farms withdrew:



**Figure 1.** Mean number of lactating cows in the 91 originally enrolled National Cohort of Dairy Farms (NCDF) herds over the 2007–2008 data collection. Herds are ordered from largest to smallest.

one due to a farm fire early in 2008, and 2 ceased dairy operations during the year. The time-weighted average number of farms participating in NCDF sampling for the 2-yr period was 83 (Table 1).

Descriptive information of NCDF herds was calculated using all of the data on hand from the 91 originally enrolled herds. The distribution of herd size was right-skewed (Figure 1), but approximately equivalent to the weighted average herd size of all DHI-recording dairy farms in corresponding regions over 2007 to 2008. Cows ranged from parities 1 to 15 (median = 2). Mean 305-d



**Figure 2.** Mean 305-d milk production (kg) of cows in the 91 originally enrolled National Cohort of Dairy Farms (NCDF) herds over the 2007–2008 data collection. Herds are ordered from highest producing to lowest producing.

**Table 2.** Average numbers of lactating cows, average daily milk production per cow per day (kg), geometric (gSCC) and production-weighted arithmetic (waSCC) mean SCC ( $\times 1,000$  cells/mL) on all 91 originally enrolled National Cohort of Dairy Farms (NCDF) herds throughout the 2007–2008 data collection

Characteristic	Mean	SD	Minimum	Median	Maximum	Region-weighted Canadian mean <sup>1</sup>
Number of lactating cows	85	48	32	69	326	77
Daily milk production/cow	32	4	21	32	47	32
gSCC	87	30	34	85	197	
waSCC	236	77	103	228	450	249

<sup>1</sup>Weighting based on Canadian averages distributed similarly to farms within NCDF (20% Alberta, 30% Ontario, 30% Québec, 20% Atlantic Canada).

milk production was 9,781 kg (range 7,734 to 12,377; SD = 859) for the 91 herds (Figure 2). Distributions of means of DHI-measured SCC by farm, by DHI test day during the data collection, were also right-skewed (Table 2).

### Sampling Compliance

Full compliance with clinical mastitis follow-up samples was 59%, with another 16% of cases having the initial sample and at least 1 follow-up sample. Technician-sampled incidence series on lactating cows achieved 89% compliance overall. Substitutions were not made for cows culled during IMI incidence sampling. Dry cow, fresh cow sampling had 62% overall compliance. A further 12% of these series were missing only 1 fresh cow sample; substitutions were made for cows culled during the dry cow, fresh cow sampling. Incomplete sets were more likely to be missing 1 (12% of total sets) or both (19%) fresh cow samples than 1 (1.1%) or both (0.47%) dry cow samples. All monthly bulk tank samples were taken for 44% of the herds, whereas another 24% had only 1 or 2 missing bulk tank samples.

For questionnaires, 99% of producers completed the first and second questionnaires, whereas 95% of producers completed the third questionnaire.

### Pathogen Recovery

Milk samples (bulk tank, quarter, and composite) numbered 134,346, and pathogen recovery was greater

than expected. Quantification of pathogen recovery is summarized in Tables 3 and 4. *Staph. aureus* was the most commonly isolated pathogen from clinical mastitis and follow-up samples, as well as the second most common isolate from all other samples. The CNS were the most commonly isolated family of pathogens from nonclinical lactating cows sampled postcalving, during the middle of lactation, and before dry-off. *Corynebacterium* spp. were also commonly isolated in all sample series besides clinical mastitis. *Klebsiella* species, *Streptococcus uberis*, *Streptococcus dysgalactiae*, *Staphylococcus hyicus*, *Streptococcus canis*, *Enterobacter*, *Nocardia*, *Prototheca*, *Arcanobacterium pyogenes*, yeasts, and fungi were also cultured, although less frequently. Isolates of *Staph. aureus* or *Strep. agalactiae* were reported for contaminated (3+ colony types) samples. Isolates recovered from samples considered to be of interest according to evaluation criteria are stored in the MPCC (n = 16,915).

### Data Entry Errors

Suspect values have been corrected through reference to original forms, and 67 problematic values that could not be resolved in this way were flagged as suspicious. A validation audit of a random sample of 750 bacteriology results forms was conducted to assess the bacteriology data entry error rate. The error rate was estimated to be 1.07% (95% confidence interval: 0.50, 2.02%). Data entry errors detected during the validation process were corrected in the database. The estimated data entry

**Table 3.** Pathogen recovery from samples from all 91 originally enrolled National Cohort of Dairy Farms (NCDF) herds over the 2007–2008 data collection period

Number of isolates recovered/sample	Clinical mastitis and follow-up	Proportion of clinical samples (%)	Lactational incidence-density samples	Proportion of lactational samples (%)	Dry cow – fresh cow samples	Proportion of dry cow – fresh cow samples (%)
0	2,031	27	37,001	43	12,741	32
1	3,528	46	27,744	32	12,593	31
2	1,040	13	10,056	12	5,936	15
3+	1,109	14	11,068	13	8,810	22
Total samples	7,622	100	85,576	100	39,742	100

**Table 4.** Most common pathogens recovered (as % of samples yielding 1 or 2 recoverable isolates) from differing sample types from all 91 originally enrolled National Cohort of Dairy Farms (NCDF) herds over the 2007–2008 data collection period

Sample type	Most common pathogen isolated		Second most common pathogen isolated		Third most common pathogen isolated	
	Species	%	Species	%	Species	%
Clinical mastitis and follow-up	<i>Staphylococcus aureus</i>	13	<i>Escherichia coli</i>	11	<i>Enterococcus</i> spp.	8.0
Lactational incidence-density samples	CNS	5.4	<i>Staph. aureus</i>	2.4	<i>Corynebacterium</i> spp.	2.2
Pre-dry-off samples	CNS	6.5	<i>Staph. aureus</i>	3.0	<i>Corynebacterium</i> spp.	2.7
Postcalving samples	CNS	4.0	<i>Staph. aureus</i>	3.0	<i>Corynebacterium</i> spp.	1.1

error rate was judged sufficiently low to necessitate double entry of all the forms.

## DISCUSSION

The research platform documented here overcomes some of the data collection hurdles associated with udder health research by gathering data in a collective effort of several different research teams. The NCDF is a nationally-relevant research population with data spanning multiple levels, linked together at each level. The data and material is available for use not only by CBMRN mastitis researchers, but also by other researchers across North America and worldwide. The strengths and valid applications of this unique collection of data are evident. These data, collected over a 2-yr time period on a national cohort of animals, offer a unique opportunity to fully investigate the effects on udder health of interactions between dairy cows, environment, and immune function. Furthermore, such longitudinal data will allow for the identification of factors associated with the occurrence of new IMI rather than just IMI prevalence: an important step in making causal inferences.

Like every large-scale field data collection, the NCDF data has some limitations. Geographic restrictions of a large country like Canada made random sampling of farms across the dairy industry impossible. Convenience samples of farms were chosen because farms needed to be close and accessible to the regional centers, as well as dedicated to active participation in the data collection. Although this is not a random sample, attempts were made to select farms that accurately represent the Canadian dairy industry. Moreover, an effort was made to prevent over-weighted selection of high-performing or low-performing herds with respect to udder health by attempting to recruit equal numbers of herds in 3 strata of a 12-mo average monthly BTSCC. Although uniform distributions among SCC strata were not achieved, the distributions achieved provide a representative character of the farms in the particular region. Selection bias may be encountered in convenience sampling, and in this instance, herds willing to participate

in this data collection may be likely to be run by more diligent producers who were willing to follow suggested protocols, notwithstanding the fact that a substantial number of herds were not high performers in terms of BTSCC. Analyses of risk factors and outcomes using this data will need to take into account this potential selection bias, as more care may have been taken in the management of some of these farms than on Canadian dairy farms as a whole.

Comparison of these herds with Canadian herds shows that the NCDF, although not inherently free from unmeasured biases, appears to be representative of Canadian dairies (Table 2). Overall, NCDF coordinating centers were established in regions that, together, comprise 93% of Canadian dairy farms enrolled in milk recording and 91% of all Canadian dairy farms (Canadian Dairy Information Centre, 2009). The NCDF restriction to herds milking mostly Holstein-Friesian cows and milking twice daily, were representative of 93% (Canadian Dairy Commission, 2009), and 96% (CanWest DHI, 2008 Progress Report) of Canadian dairy herds, respectively. The average number of lactating cows on NCDF farms was 85, comparable to a region-weighted average in commercial dairy farms across Canada of 77 (CanWest DHI, 2008 Progress Report). Housing systems were comparable between the NCDF farms and the Canadian industry. Nationally, a region-weighted average of 36% of farms house cows in freestalls (Olde Riekerink et al., 2010); the NCDF had 34% of farms housing in freestalls. The region-weighted average of milk production per Holstein cow reported by CanWest DHI across 2007 and 2008 was 9,784 kg of milk produced per cow (2008), and NCDF farms averaged 9,876 kg per cow. Clinical mastitis on Canadian dairy farms has been estimated at an incidence of 0.22 cases per cow-yr (Olde Riekerink et al., 2008), and preliminary statistics for the NCDF show farms achieving an incidence of 0.20 cases per cow-yr.

Due to shipping distances to regional bacteriology laboratories and the need to accommodate the laboratories' weekly sample analysis capacity, freezing of milk samples was necessary. Freezing, however, is not without detriment from a bacteriological perspective.



Schukken et al. (1989) reported that freezing significantly decreased positive cultures of *Escherichia coli* and *A. pyogenes* in clinical mastitis cases by approximately 20%, and extending the length of time frozen further contributed to this decrease. Biddle et al. (2004) also reported a continuous decline in the log cfu/mL of *Mycoplasma* spp., recovered as samples that were stored frozen for longer periods, with as much as 45% loss in the number of samples culturing positive for *Mycoplasma* spp. Because of this low recovery rate from frozen samples and its low prevalence in Canadian herds (Olde Riekerink et al., 2006), *Mycoplasma* spp. identification was not included in the bacteriological culture protocol. Conversely, Schukken et al. (1989) showed that freezing significantly increased the number of CNS-positive cultures. These changes also became more prominent as length of freezing was extended. Schukken et al. (1989) also observed no significant effect of freezing or length of freezing on the ability to culture streptococci or *Staph. aureus*. This is in agreement with more recent work, as Godden et al. (2002) found no significant effect of freezing on the mean category score for cfu/mL of *Staph. aureus*, although Sol et al. (2002) demonstrated a higher isolation percentage of *Staph. aureus* in frozen samples. Murdough et al. (1996) also found no difference in the viability of *Staph. aureus*, 3 species of CNS, *Strep. agalactiae*, *Strep. dysgalactiae*, *Strep. uberis*, *Corynebacterium bovis* or *E. coli* in quarter milk samples frozen for 6 wk. These contrasting results with regard to CNS, *E. coli*, and *Staph. aureus* must be taken into account when utilizing the data presented here. They will also be of interest to researchers performing bacteriological analyses on frozen samples, especially when freezing duration is extended. Contrasting results between researchers suggests that factors other than simply frozen storage time are operating to influence apparent associations between duration of freezing and species-specific recovery of pathogens from milk samples.

Previous studies have shown that cryostorage times of up to 28 d produce a small measurable change in SCC in frozen versus fresh milk samples (Barkema et al., 1997; Trenholm et al., 2007). In pilot studies, comparison between fresh samples, frozen samples, refrozen and preserved samples, refrozen-preserved and stored samples, and refrigerated samples showed high concordance (as described in Lin, 1989) of SCC (all >0.9), but decrease in concordance correlation appreciated over time (C. D. Calloway, I. R. Dohoo, and G. P. Keefe, Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, Canada, personal communication). These high-concordance correlations show that a frozen-thawed-preserved sample can be used to accurately measure SCC using

fluoro-opto technology, although results may be less accurate when flow cytometry is used. C. D. Calloway, I. R. Dohoo, and G. P. Keefe (Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, personal communication) also found that samples were viable (concordance correlation: 0.88, 95% confidence interval: 0.85, 0.91) when frozen for 7 d, thawed, preserved, refrozen, stored on ice for 3 d and refrigerated for up to 1 wk.

Data quality is a very important issue in field data collection, comprising completeness as well as accuracy. In the NCDF, we broadly encouraged completeness and accuracy of sampling and supplying of information in several ways: assuring confidentiality to producers, doing our best to convey information (mainly bacteriology results) back to producers, promoting a sense of participation and ownership among producers and technicians through training and incentives, and on-going data validation. Technicians served as the main point of contact to producers. Technician retention was 100%, permitting them to earn producer confidence and to obtain a large amount of information from producers. All parties involved could see the results of their efforts and the scientific value of the data, which served to keep them motivated to collect quality data.

A very substantial amount of material and linked data was collected from the NCDF. These are useful for studying many different questions pertinent to clinical mastitis in developed dairy industries. Nevertheless, some data may be incomplete. Producer feedback indicated some misunderstanding as to what defines a clinical mastitis case, although producers had been clearly instructed to collect samples from any quarter with abnormal milk, regardless of intention to treat. Some considered a few clots at the beginning of milking not sufficient to qualify as abnormal, whereas others mistakenly collected samples from cows with high SCC and no abnormal milk. Producers were also less likely to sample or resample clinical cases from cows they already believed to have chronic infections. Thus, the number of clinical mastitis cases may be underestimated. Similarly, although clinical samples were always prioritized in the labs, the realized interval from sampling date to the date results were reported was sometimes long, possibly diminishing the producer-perceived value of clinical case sampling. However, when producer-reported records of clinical mastitis were compared with the number of samples submitted over the 2 yr, producers collected samples for 6 more cases of clinical mastitis on average than they reported in their records ( $P = 0.06$ ; S. Dufour, unpublished data).

The sample series taken for IMI incidence estimation from lactating cows were the most complete, as they were taken by regional technicians according to

a pre-planned schedule. Missed samples were typically due to cow removal during the 6-wk sampling interval. For pre-dry and postcalving series, farmers tended to collect the first samples (either pre-dry or postcalving) most often, with follow-up sample collection being less complete. Some, but not all, of decreased completion was again due to culling or death of enrolled cows. Logistical difficulty also exists in incorporating sample-taking into already existing farm routines. These deficits in sample series completion may lead to less precise estimates of associations with IMI, but not necessarily less valid estimates. Loss of information from incomplete sampling series would lead to research bias only if the reasons for incomplete sampling are related to the association being estimated. Future cohort data collection undertakings should seek ways to optimize technicians' roles in promoting more feasible sampling protocol completion.

It may be speculated that contamination results of the samples taken in this data collection platform are multifactorial. Although every attempt was made to ensure good sampling practices, having several technicians working at different sites and on several farms and training farmers on a set protocol for milk sampling is a monumental task. As the data collection progressed, laboratories reported results of bacteriologic cultures, and technicians were often vigilant to recognize farms where the proportions of contaminated samples seemed excessive. These farms were pinpointed for review of sampling technique.

Another challenge was the requirement of collection of at least 30 mL of milk for accurate SCC measurement. Due to cow movement, collection of this amount of milk into a single sample vial in an aseptic manner is difficult and may have led to increased frequency of contamination in these sampling series. On the other hand, using a single sample vial for both bacteriological culture and SCC simplified the sampling protocols for farm personnel, and reduced the risks of logistical failure of sample storage and shipping procedures. This latter risk in a high-volume, geographically extensive, and extended-time cohort study should not be underestimated.

In these data, the lowest contamination percentages were found in lactational samples (13%; mostly technician-collected), with higher percentages in clinical samples (14%) and samples taken pre-dry-off and postcalving (22%; mostly farmer-collected). Clinical samples were perhaps easiest to collect because only one quarter was sampled at each time. Alternatively, pre-dry-off quarter and composite samples were taken at a time when cows potentially had less milk, making sampling difficult. Similarly, cows postcalving tend to be less tolerant of sampling, potentially leading to

higher contamination percentages in quarter and composite samples taken in these series.

Laboratories in the MLN also used strict definitions of contamination: when 3 or more dissimilar colony morphologies were cultured, the sample was classified as contaminated. This strict adherence to the definitions set forth allowed comparability between labs, but may have also increased contamination percentages.

Researchers using this data should be aware that the effect on data validity of contaminated samples, specifically through missing data bias, will depend on methods of analysis. If samples which should have been identified with either major or minor pathogens or should have been deemed "no growth" were instead classified as contaminated, diagnosis with an IMI may have been missed. This may decrease the power of studies performed on these data. There is no reason, however, to believe that noninfected cows would have had more or less contaminated samples than did infected cows, which, in turn, leads to nondifferential misclassification of IMI. However, bias due to misclassification will need to be dealt with by individual researchers using these data.

Any data collection effort that relies on farm personnel executing sampling protocols, as is virtually necessitated by large intensive studies, must confront issues of sampling compliance. It was noted in this data collection that the worst compliance concerned those protocols that were most inconsistent with farmers' existing work patterns. Future longitudinal studies on the scale of the present study should explore ways of optimizing the balance between the costs of project technicians collecting samples on-farm and compliance to farmer-executed sampling protocols. Several methods of improving compliance in milk sampling or producer response rate to questionnaires and health record-keeping have been suggested. The effect of monetary incentives for complete sets of samples was observed, as the on-farm culture project did offer producers CA\$30 per completed record and follow-up samples for each clinical mastitis case. Compliance with the complete data collection during the on-farm culture project was 75%, compared with 59% in the overall CBMRN clinical mastitis sampling series (K. A. MacDonald, Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, personal communication). Although incentives were regularly given to participating producers throughout the NCDF data collection to boost morale, an incentive program related to completion for all sampling schemes could have been used to augment the data collection, but would also serve to dramatically increase the overall budget and might potentially lead to over-sampling. Good communication in the form of results reporting, producer

meetings involving discussion, and distribution of comparative herd reports were employed, although rapid results reporting accompanying incentives might have had a more lasting effect on producers' prioritization of sample-taking in their daily activities.

Pathogen recovery was greater than had been originally forecast, especially for minor pathogens such as CNS and *Corynebacterium* spp. Pathogen isolates in the MPCC are available for further study by collaborating researchers upon request.

Data from the NCDF supports research in several areas related to udder health and can be applied to a broad array of topics still needing further investigation. This data collection offers a unique research resource and presents opportunities for additional collaborative research.

### ACKNOWLEDGMENTS

The authors thank all of the regional technicians (Theresa Andrews, Meliza Morris, Natasha Robinson and Adele Veinot, University of Prince Edward Island, Canada; François Dubois, Université de Montréal, Quebec, Canada; Michael MacLean, University of Guelph, Kemptville Campus, Ontario, Canada; Andrea Wascko and Anke Wellen, University of Calgary, Alberta, Canada), laboratory staff (Serge Messier, Université de Montréal; Musangu Ngeleka and Maggie Schwab, University of Saskatoon, Saskatchewan, Canada; Doris Poole and Shana Richard, University of Prince Edward Island) and students that assisted with this data collection. Special thanks are reserved for all participating producers for asking valuable questions and continually propelling us forward. Ken Leslie and David Kelton were also instrumental in providing a connection with Kemptville Campus of the University of Guelph which served as the Ontario regional centre. This research was financed by the National Sciences and Engineering Research Council of Canada, Alberta Milk, Dairy Farmers of New Brunswick, Nova Scotia, Ontario and Prince Edward Island, Novalait Inc., Dairy Farmers of Canada, Canadian Dairy Network, Agriculture and Agri-Foods Canada, Public Health Agency of Canada, Technology PEI Inc., Université de Montréal, and the University of Prince Edward Island, through the Canadian Bovine Mastitis Research Network.

### REFERENCES

Barkema, H. W., Y. H. Schukken, T. G. J. M. Lam, M. L. Beiboer, H. Wilmink, G. Benedictus, and A. Brand. 1998. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. *J. Dairy Sci.* 81:411-419.

Barkema, H. W., J. Van der Schans, Y. H. Schukken, A. L. W. De Gee, T. J. G. M. Lam, and G. Benedictus. 1997. Effect of freezing on somatic cell count of quarter milk samples as determined by a Fossomatic electronic cell counter. *J. Dairy Sci.* 80:422-426.

Biddle, M. K., L. K. Fox, D. D. Hancock, C. T. Gaskins, and M. A. Evans. 2004. Effects of storage time and thawing methods on the recovery of *Mycoplasma* species in milk samples from cows with intramammary infections. *J. Dairy Sci.* 87:933-936.

Bradley, A. J., K. A. Leach, J. E. Breen, L. E. Green, and M. J. Green. 2007. Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales. *Vet. Rec.* 160:253-258.

Browning, J. W., G. A. Mein, P. Brightling, T. J. Nicholls, and M. Barton. 1994. Strategies for mastitis control: Dry cow therapy and culling. *Aust. Vet. J.* 71:179-181.

Canadian Dairy Commission. 2009. Canadian Dairy Commission. Accessed May 19, 2009. <http://www.cdc-ccl.gc.ca/cdc/index-eng.php>.

Canadian Dairy Information Centre. 2009. Enrollments on milk recording. Accessed May 19, 2009. [http://www.dairyinfo.gc.ca/index\\_e.php?s1=df-fcil&s2=mrr-pcle&s3=dhi-agbl&page=emr-ipcl](http://www.dairyinfo.gc.ca/index_e.php?s1=df-fcil&s2=mrr-pcle&s3=dhi-agbl&page=emr-ipcl).

CanWest DHI. 2008. Progress report statistics. Accessed May 19, 2009. [http://www.canwestdhi.com/pdf\\_files/2008%20progress%20report%20statistics.pdf](http://www.canwestdhi.com/pdf_files/2008%20progress%20report%20statistics.pdf).

Dufour, S., H. W. Barkema, L. DesCôteaux, T. DeVries, I. R. Dohoo, K. K. Reyher, J.-P. Roy, and D. T. Scholl. 2010. Development and validation of a bilingual questionnaire for measuring udder health related management practices on dairy farms. *Prev. Vet. Med.* 95:74-85.

Funk, D. A., A. E. Freeman, and P. J. Berger. 1982. Environmental and physiological factors affecting mastitis at drying off and post-calving. *J. Dairy Sci.* 65:1258-1268.

Godden, S. M., J. T. Jansen, K. E. Leslie, N. L. Smart, and D. F. Kelton. 2002. The effect of sampling time and sample handling on the detection of *Staphylococcus aureus* in milk from quarters with subclinical mastitis. *Can. Vet. J.* 43:38-42.

Halasa, T., K. Huijps, O. Østerås, and H. Hogeveen. 2007. Economic effects of bovine mastitis and mastitis management: A review. *Vet. Q.* 29:18-31.

Hogan, J. S., R. N. Gonzalez, R. J. Harmon, S. C. Nickerson, S. P. Oliver, J. W. Pankey, and K. L. Smith. 1999. Laboratory Handbook on Bovine Mastitis. 1st ed. NMC, Inc., Madison, WI.

Lin, L. I. 1989. A concordance coefficient to evaluate reproducibility. *Biometrics* 45:255-268.

Murdough, P. A., K. E. Deitz, and J. W. Pankey. 1996. Effects of freezing on the viability of nine pathogens from quarters with subclinical mastitis. *J. Dairy Sci.* 79:334-336.

National Cohort of Dairy Farms (NCDF). 2009. National Cohort of Dairy Farms reference manual. Accessed May 19, 2009. <http://www.mastitisnetwork.org>.

Neijenhuis, F., H. W. Barkema, H. Hogeveen, and J. Noordhuizen. 2000. Classification and longitudinal examination of callused teat ends in dairy cows. *J. Dairy Sci.* 83:2795-2804.

Olde Riekerink, R. G. M., H. W. Barkema, D. F. Kelton, and D. T. Scholl. 2008. Incidence rate of clinical mastitis on Canadian dairy farms. *J. Dairy Sci.* 91:1366-1377.

Olde Riekerink, R. G. M., H. W. Barkema, D. T. Scholl, D. E. Poole, and D. F. Kelton. 2010. Management practices associated with the bulk-milk prevalence of *Staphylococcus aureus* in Canadian dairy farms. *Prev. Vet. Med.* 97:20-28.

Olde Riekerink, R. G. M., H. W. Barkema, S. Veenstra, D. E. Poole, R. T. Dingwell, and G. P. Keefe. 2006. Prevalence of contagious mastitis pathogens in bulk tank milk in Prince Edward Island. *Can. Vet. J.* 47:567-572.

Rothman, K. J., S. Greenland, and T. L. Lash. 2008. Modern Epidemiology. 3rd ed. Page 97. Lippincott Williams and Wilkins, Philadelphia, PA.

Schukken, Y. H., J. A. H. Smit, F. J. Grommers, D. Vandegeer, and A. Brand. 1989. Effects of freezing on bacteriologic culturing of mastitis milk samples. *J. Dairy Sci.* 72:1900-1906.

- Sears, P. M., and K. K. McCarthy. 2003. Diagnosis of mastitis for therapy decisions. *Vet. Clin. North Am. Food Anim. Pract.* 19:93–108.
- Seegers, H., C. Fourichon, and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. *Vet. Res.* 34:475–491.
- Sol, J., O. C. Sampimon, E. Hartman, and H. W. Barkema. 2002. Effect of preculture freezing and incubation on bacteriological isolation from subclinical mastitis samples. *Vet. Microbiol.* 85:241–249.
- Trenholm, M. J., G. P. Keefe, and I. R. Dohoo. 2007. Effect of freezing of quarter milk samples on somatic cell count measured using flow cytometry and flouro-opto-electronic methods. Pages 234–235 in *NMC Ann. Mtg. Proc.*, Harrisburg, PA. NMC, Inc., Madison, WI.
- Zadoks, R. N., H. G. Allore, H. W. Barkema, O. C. Sampimon, G. J. Wellenberg, Y. T. Gröhn, and Y. H. Schukken. 2001. Cow- and quarter-level risk factors for *Streptococcus uberis* and *Staphylococcus aureus* mastitis. *J. Dairy Sci.* 84:2649–2663.