Rare Diseases 1, e25212; 2013; © 2013 Landes Bioscience

Lifetime risk estimators in epidemiological studies of Krabbe Disease

Review and Monte Carlo comparison

Alexander H. Foss,[1] Patricia K. Duffner[2] and Randy L. Carter[1,*]

[1]Department of Biostatistics; Population Health Observatory; School of Public Health and Health Professions; University at Buffalo; State University of New York; Buffalo, NY USA; [2]Hunter James Kelly Research Institute; Department of Neurology; University at Buffalo; State University of New York; Buffalo, NY USA

Keywords: epidemiology, lifetime risk, incidence, leukodystrophy, globoid cell, lysosomal storage diseases, metabolism, inborn errors, Monte Carlo method, prevalence

Abbreviations: GALC, galactosylceramidase; nmol/h/mg, nanomol per hour per milligram protein; Dx method, method of calculating lifetime risk defined in Table 2; DOB method, method of calculating lifetime risk defined in Table 2; LT method, method of calculating lifetime risk defined in Table 2; MC, Monte Carlo

This review addresses difficulties arising in estimating epidemiological parameters of leukodystrophies and lysosomal storage disorders, with special focus on Krabbe disease. Although multiple epidemiological studies of Krabbe disease have been published, these studies are difficult to reconcile since they have used different study populations and varying methods of calculation. Confusion exists regarding which epidemiological parameters have been estimated; the current review shows that most previous estimates can be properly interpreted as lifetime risk at birth. One of the most common estimation methods is shown to be inaccurate, while two other methods are shown to be approximately accurate. Based on the results of the current paper, recommendations are made that are expected to improve the quality of future studies of Krabbe disease. It is anticipated that these recommendations will be applicable to epidemiological studies of other lysosomal storage disorders, as well as any other rare diseases diagnosed with enzymatic screening.

Introduction

Krabbe disease (OMIM # 245200) arises as a result of deficient galactosylceramidase (GALC; E.C. 3.2.1.46) activity, which is caused by a mutated *GALC* gene. GALC is a degradative lysosomal enzyme, and thus Krabbe disease is classified as a lysosomal storage disorder[1,2] in addition to its classification as a leukodystrophy. Since the *GALC* gene was cloned,[3] about 100 mutations to the *GALC* gene have been identified.[4] Some mutations are associated with well-specified phenotypes. For example, individuals homozygous for the 30-kb deletion (OMIM # 606890.0002) generally develop the early infantile form of the disease. Other mutations, however, lead to substantial phenotypic variability, both intrafamiliar and otherwise.[5]

Krabbe disease is relatively difficult to diagnose due to its rarity and phenotypic heterogeneity. Primary care physicians may never see a case in their career. The proportion of Krabbe patients that never receive a correct diagnosis is unknown. Diagnostic accuracy is likely to vary drastically based on a family's access to adequate medical care and physician experience. In an international survey of 334 families self-reported to have a member suffering from Krabbe disease, 103 reported the length of time between initial symptom onset and correct diagnosis. Mean time to diagnosis was 5.3 mo,[6] suggesting a substantial difficulty in the diagnosis of Krabbe disease. It has been suggested that individuals with adolescent and adult forms are particularly susceptible to misdiagnosis due to reduced awareness of later-onset subtypes.[6,7] While symptom onset during infancy is generally associated with a more predictable and fulminant clinical course, the course of patients with later onset is quite variable,[6,8,9] and this variability is likely to contribute to diagnostic challenges facing practitioners.

Enzymatic screening and mutation analysis have improved antemortem diagnostic capabilities, offering good sensitivity in individuals with putative symptoms of Krabbe disease. As might be expected, challenges arise when these methods are applied to a large population of asymptomatic individuals. In 2006, New York State began screening all newborns for Krabbe disease, and by 2008 roughly 550,000 newborns had been screened. Four newborns were placed in the high-risk category based on GALC activity below 0.15 nmol/h/mg; two of these newborns received a positive diagnosis based on subsequent mutation analysis and abnormal neurodiagnostic studies. The other two had negative neurodiagnostic studies and remained asymptomatic at follow-up at 8 and 16 mo of age. Six were placed in the moderate-risk

*Correspondence to: Randy L. Carter; Email: rcarter@buffalo.edu
Submitted: 03/03/13; Revised: 05/14/13; Accepted: 05/29/13;
Published Online: 06/18/13
Citation: Foss AH, Duffner PK, Carter RL. Lifetime risk estimators in epidemiological studies of Krabbe Disease: Review and Monte Carlo comparison. Rare Diseases 2013; 1:e25212; http://dx.doi.org/10.4161/rdis.25212

category (GALC between 0.16–0.29 nmol/h/mg protein) and 15 were placed in the low-risk category (0.3–0.5 nmol/h/mg protein). All moderate- and low-risk children were asymptomatic at follow-up.[5] Presently, only subsequent follow-up will be able to determine whether these asymptomatic at-risk children will develop symptoms at a later date, whether they are carriers of the disease or whether they simply harbor non-pathogenic GALC-lowering polymorphisms.[9]

Epidemiological analysis must also consider missed diagnoses due to censored cases: individuals who would have developed symptoms of Krabbe disease but died of unrelated causes before symptom onset. The impact of censoring on measures of lifetime risk at birth will be quantified in subsequent sections of the current paper.

Previous Epidemiological Studies of Krabbe Disease

The literature on the epidemiology of Krabbe disease was surveyed, and any article that reported an estimated rate of disease (prevalence, incidence, birth prevalence, etc.) was considered for review.[33-36] These previous epidemiological studies of Krabbe disease are summarized in **Table 1**. There are numerous published studies that reported raw numbers of patients suffering from Krabbe disease in a given population.[10-15] Since these studies did not estimate disease rates or proportions relative to a reference population, they will not be reviewed in the current paper. Seven studies aimed to estimate the proportion of individuals who will eventually be diagnosed with the disease in a given birth cohort,[7,16-21] which, as shown below, is best referred to as lifetime risk at birth. Three studies estimated lifetime risk for other disease categories, but also provided sufficient information to estimate it for Krabbe disease.[22-24] Three studies estimated this proportion for isolated sub-populations within Israel and will not be reviewed further.[25-27] Several book chapters provided epidemiological estimates without describing the methods used for estimation.[28-31]

The studies reporting the proportion of individuals affected by Krabbe disease all generally sought to approximate the same quantity: the proportion of newborns that will develop the disease at any point in their life. It will be asserted below that this quantity corresponds best to lifetime risk at birth. Reported calculations typically involved dividing the number of diagnoses in a given time period by the number of births within a corresponding time period. A broad and often conflicting variety of terms were used to describe this quantity in the literature. Terms used include incidence (e.g., ref. 16), prevalence (e.g., ref. 18) and birth prevalence (e.g., ref. 20). Meikle et al.[18] used the term prevalence to denote calculations including pre- and postnatal diagnoses and incidence to denote calculations that include only postnatal cases.

The usage of the terms above to describe the proportion in question does not match formal definitions in epidemiological publications. This misuse of terminology in the literature presents two problems that will subsequently be addressed: first, identifying what parameter these studies intended to estimate and second, evaluating the accuracy of those estimations.

What parameters are estimated? Given common incorrect use of terms used to describe these epidemiological quantities, the correct interpretations of reported rates in the seven studies listed above[7,16-21] must be inferred from their stated methods. The proportion of newborns that are, or will be, affected by a disease is best described as birth prevalence or lifetime risk at birth, depending on the nature of the disease being studied. For conditions that are exclusively incident at or near birth, such as cleft palate, they are equivalent, since the number of affected newborns is equal to the number of individuals who will develop the disease at any point in their lives. For conditions with a variable age of onset, however, birth prevalence refers only to those who have the disease at birth. Since the majority of studies included juvenile or adult onset Krabbe cases,[7,17-20,22-24] birth prevalence does not adequately describe the estimates in question. It is conceivable that the studies could be describing the birth prevalence of necessary and sufficient factors for developing the disease later in life, but it is not fully known what these factors are (as illustrated above by the challenges of newborn screening and unknown predictive validity of many pathogenic alleles). We propose that lifetime risk at birth describes the desired proportion more effectively. Lifetime risk is a special case of cumulative incidence (which we distinguish from incidence *rate*) in which the period of time studied is the entire remaining lifetime. Lifetime risk can be estimated conditional on disease-free survival up to a specified age, and thus lifetime risk at birth should be specified as such. For discussions of all the above epidemiological terms, readers are referred to several resources listed in the bibliography.[32-37]

The seven studies listed above do not capture information regarding which age groups are at greater risk of death due to Krabbe disease. Only one study[38] aimed to measure age-specific mortality rates. Using national and state death certificates in the United States, the number of deaths of individuals with Krabbe disease was estimated for the years 1999–2004. Mortality rates were estimated separately for children less than five years of age and compared with mortality rates of individuals greater than or equal to five. It was found that mortality in children less than five years of age (0.994 deaths per year per million children) was greater than mortality in older individuals (0.007 deaths per year per million individuals). Barczykowski et al.[38] also showed that the yearly mortality rate in children less than 5 y of age due to Krabbe disease is a good approximation to the incidence rate of the early infantile form of Krabbe disease.

It is not possible to evaluate the estimates provided in any available book chapters,[28-31] since the methods used were not described. These estimates are referred to as incidence, although estimates are not given in units involving time, so they cannot be incidence rates. It is presumed that these estimates correspond to lifetime risk at birth, as described above, but without a description of the calculations used this cannot be determined.

How accurate are the estimates? Although it can be assumed that the seven studies in question[7,16-21] produced ratios that can be interpreted as estimates of lifetime risk from birth, it is not immediately clear that they succeeded in measuring this quantity accurately. These studies did not follow cohorts longitudinally, but

Table 1. Previous studies of Krabbe epidemiology

Reference	Methods described?	Study population	Reported measure	Numerator	Denominator	Estimate, per 1 million units
Hagberg et al. 1969[16]	Yes	Sweden, 1953–1967	incidence	32 postnatal cases born 1953–1967	1.66 million births 1953–1967	19
Heim et al. 1997[17]	Yes; DOB	Germany, 1981–1989	incidence ◇	34 cases born 1981 – 1989 and diagnosed before 1994	All births 1981–1989	6.0
Meikle et al. 1999[18]	Yes; Dx	Australia, 1980–1996	prevalence	30 pre- and postnatal diagnoses 1980–1996	4,222,323 births 1980–1996	7.1
Meikle et al. 1999[18]	Yes; Dx	Australia, 1980–1996	incidence	21 postnatal diagnoses 1980–1996	4,222,323 births 1980–1996	5.0
Poorthuis et al. 1999[20]	Yes; DOB	Netherlands, 1971–1995	birth prevalence	63 pre- and postnatal diagnoses 1971–1996	4,677,849 births 1971–1995	13.5
Ozkara and Topcu 2004[21]	Yes	Turkey, 1997–2002, age < 5	incidence	65 postnatal cases diagnosed 1997–2002, born 1997–2002	6,500,000 live births between 1997–2002	10
Pinto et al. 2004[19]	Yes; DOB	North Portugal, 1984–1999	birth prevalence	9 pre- and postnatal diagnoses 1988–2001	All births 1984–1999	12.1
Poupetova et al. 2010[7]	Yes; DOB	Czech Republic, 1975–2008	birth prevalence	13 cases born 1977–2002, diagnosed 1975–2008	3,249,150 live births, 1977–2002	4.0 (2.1 - 6.8)§
Barczykowski et al. 2012[38]	Yes	United States, 1999–2004, age < 5	mortality rate	19.40 expected deaths per year due to Krabbe disease	19,521,730 average population size, age < 5	0.994
Barczykowski et al. 2012[38]	Yes	United States, 1999–2004, age ≥ 5	mortality rate	1.87 expected deaths per year due to Krabbe disease	266,827,247 average population size, age ≥ 5	0.007
Zlotogora et al. 1985[25]	Yes	Druze in Israel, 1969–1983	incidence	12 postnatal cases born 1969–1983	2000 births 1969–1983	6000
Zlotogora et al. 1991[26]	Yes	Jews in Israel, 1973–1987	incidence	0 cases 1969–1983	1,000,000 births 1969–1983	0
Zlotogora et al. 1991[26]	Yes	Muslim Arabs in Jerusalem, 1973–1987	incidence	6 postnatal cases born 1979–1987	3731 births 1973–1987	1610
Macarov et al. 2011[27]	Yes	Muslim Arabs in Jerusalem, 1999–2002	incidence	6 postnatal cases born 1999–2002	3600 live births 1999–2002	1670
Macarov et al. 2011[27]	Yes	Muslim Arabs in Jerusalem, 2003–2007	incidence	4 postnatal cases born 2003–2007	4876 live births 2003–2007	820
Suzuki and Suzuki 1983[28]	No	Japan, 8 y period	incidence	17 cases over an 8 y period	Births	2–3
Suzuki et al. 1995[29]	No	Japan, 1972–1986	incidence	25 cases	Births	5–10
Suzuki et al. 1995[*,29]	No	France	incidence	Cases of Krabbe disease	Live births	6.7
Wenger et al. 2001[30]	No	United States	incidence	Cases of Krabbe disease	Births	10
Pastores and Kolodny 2006[31]	No	"General population"	incidence	Cases of Krabbe disease	No description	5.0
Applegarth et al. 2000[22]	NA †	British Columbia, 1972–1996	Lifetime risk at birth †	2 postnatal, 1 prenatal cases born 1972–1996, diagnosed 1972–1997	1,035,816 live births 1972–1996	2.9 †
Dionisi-Vici et al. 2002[23]	NA †	Italy, 1985–1997, < 18 y old	Lifetime risk at birth †	36 postnatal cases diagnosed 1985–1997	7,173,959 births 1985–1997	5.0 †
Bonkowsky et al. 2010[24]	NA †	Utah, 1999–2007, < 18 y old	Lifetime risk at birth †	2 postnatal diagnoses 1999–2007	451,171 live births, Utah, 1999–2007 ‡	4.4 †

As discussed in the current paper, many of these estimates are inaccurate or uninterpretable. *Vanier MT, cited in Suzuki et al.[29] as personal communication; ◇Heim et al.[17] reported incidence but cited the method of Claussen et al.,[45] whose method is described therein as cumulative incidence; †calculated from reported data by current author A.F., using the Dx method; note that original authors may not have intended data to be used in this manner; §95% Poisson confidence interval; ‡ Houston et al.[46]

Table 2. Definitions of key terms and the three estimation methods evaluated in the Monte Carlo (MC) simulations

Term	Definition
Diagnosis Period	The period of time in which dates of Krabbe diagnoses were collected.
Observed Case	An individual receiving a Krabbe diagnosis during the diagnosis period.
Censored Case	An individual who would have been an observed case, but died of unrelated causes before symptom onset.
Birth Period	The period of time between the earliest and latest birthdate of all observed cases.
Dx Method	Method of estimating lifetime risk at birth calculated by taking the number of observed cases divided by the number of total births during the diagnosis period. Although it is slightly more biased than the LT method, it is often the best possible method for studies of Krabbe disease.
DOB Method	Method of estimating lifetime risk at birth calculated by taking the number of observed cases divided by the number of total births during the birth period. A biased method that should not be used.
Life-Table (LT) Method	Method of estimating lifetime risk at birth that adjusts for censored cases. For a full description see, for example, Beiser et al.[32] The best method reviewed, but often unusable for studies of Krabbe disease due to the unavailability of appropriate data.

rather obtained the dates of all Krabbe diagnoses in a given time period (hereafter referred to as diagnosis period). Two distinct methods were used to calculate the denominator of the ratios reported in these papers. As will be shown below, one method is flawed, while the other yields approximately accurate estimates of lifetime risk.

Meikle et al.[18] estimated lifetime risk at birth using a method referred to in the current study as the "Dx" method (see **Table 2**). In this study, it was explicitly assumed that the number of diagnoses in the diagnosis period was equal to the number of affected births within that same diagnosis period (affected births referring to any individual who would ever develop symptoms) and that the appropriate denominator must be the number of births occurring in the diagnosis period. Assuming that the effect of censoring is minimal, a constant birth rate, a constant diagnosis rate and that all affected individuals receive a diagnosis, this equivalence must hold true, as shown in **Figure S1**. Therefore, given these assumptions, the Dx method appears to be an approximately accurate means of estimating the lifetime risk of the disease.

Four studies[7,17,19,20] estimated lifetime risk at birth using a method referred to in the current study as the "DOB" method (see **Table 2**). In these studies, the earliest and latest birth date of cases diagnosed in the diagnosis period was determined. The total number of births in the population within this birth range was used as the denominator. While this method ensures that all potentially relevant births are reflected in the denominator, it introduces a serious source of error. Any case of Krabbe disease that was born within the birth range and died before the beginning of the diagnosis period remains uncounted in the numerator, resulting in underestimated lifetime risk. Furthermore, the beginning and end of the birth period is determined by the two most extreme birthdates. Since the sample range of birthdates increases with increasing sample size, the severity of this error increases with increased sample size. This source of error was recognized by some study authors, who removed cases that they believed would excessively inflate the denominator.[7,17,20] None of these authors reported a quantitative method of selecting cases for removal, and thus their method is not reproducible by a third party. It appears that most of the cases were removed due to an early birth year, which, had they not been removed, would have greatly increased the number of births included in the denominator.

The accuracy of the Dx and DOB methods, and the influence of censored cases on the Dx method, is further analyzed in a Monte Carlo simulation study that is described in the next section of the current paper.

Ozkara and Topcu[21] studied the occurrence of Krabbe disease in children 0–5 y of age during the five year period from 1997–2002 in Turkey. They included for analysis all Krabbe patients both born and diagnosed within 1997–2002 and divided this number by the number of total births observed within 1997–2002. This method is not accurate, as the birth cohorts in the study period are not followed for equal amounts of time.

Hagberg et al.[16] studied cases of Krabbe disease that had been born from 1953–1967. They divided the total number of affected cases born during this time period by the total number of births during this time period. In this particular sample, however, the patients with the earliest and latest birthdates died in the same year they were born. In this special case, the method used by Hagberg et al.[16] is equivalent to the Dx method; thus, the conclusions reported herein regarding the Dx method are applicable to Hagberg's method.

Case ascertainment. In addition to the various ratios that have been reported and the various calculations used to approximate lifetime risk at birth, varying methods of case ascertainment have been used in the epidemiological studies reviewed above. These methodological differences are large enough to render direct comparisons across studies difficult, if not impossible. Most studies included patients of any age, while others included restricted age ranges.[23,24] Some studies included fetuses diagnosed prenatally in the analysis (e.g., ref. 19), whereas others do not appear to have done so (e.g., ref. 16). Studies report different coverage of the observed population; e.g., Heim et al.[17] reported a response rate less than 80% in the queried medical institutions, whereas Hagberg et al.[16] reported a 100% response rate). Additionally, there is likely to be a non-negligible number of undiagnosed patients with Krabbe disease in any population; the rate of undiagnosed cases is likely to vary with access to adequate primary health care providers, access to specialists in pediatric neurology and physician experience. Within any study, these factors are largely unknown, rendering corrections for these

sources of error impossible. Naturally, the measured proportions are also affected by the true disease proportion in the study population. Unfortunately, it is impossible to separate these last two sources of between-study variability, and it remains possible that country-level differences are merely attributable to country-specific rates of missed diagnoses of Krabbe disease.

Monte Carlo Evaluation of Methods for Lifetime Risk Estimation

A Monte Carlo (MC) simulation study was implemented to further evaluate the accuracy of two methods commonly used to estimate lifetime risk at birth of Krabbe disease: the Dx method and the DOB method (see **Table 2**). A third method which we shall refer to as the life-table (LT) method[32] was also evaluated. The latter method has the advantage of correcting for error due to censored cases and yielded the most accurate estimates. Although the LT method is often not practical for studies of Krabbe disease due to the extensive data required, we have included it in the simulation to serve as a reference to which the other methods can be compared.

A MC study is a computer simulation in which random samples are repeatedly generated in order to study a phenomenon of interest.[39] MC simulations have been used to investigate difficult problems across a broad range of disciplines, e.g., evolutionary biology,[40] radiology,[41] and finance.[42] In the current study, samples of populations with a known lifetime risk at birth of Krabbe disease were generated. Since the lifetime risk was known, the accuracy of the estimation methods could be calculated directly. Because the accuracy of the estimation methods were affected by fixed population parameters, these parameters were systematically varied to simulate the range of populations studied in previous studies of Krabbe epidemiology. All simulations were implemented using the R programming language.[43]

Parameters and distributional assumptions of MC simulation. The four input parameters to the (MC) study were: *T*, the length of the diagnosis window in which cases could be registered; *N*, the constant number of total births each year (both Krabbe and non-Krabbe cases); *LR*, the true lifetime risk at birth of the disease; and *A*, the average age at diagnosis of the disease. It was assumed that the birth, diagnosis and age-specific death rates did not change with calendar time and that there was no migration into or out of the observed population. It is important to note that a constant birth rate guarantees that the Dx method performs optimally; any violation of this assumption would result in biased estimates. All settings of the parameters are listed in **Table 3**. 500 MC samples were generated for each of 4320 unique combinations of parameters.

The birthdate and diagnosis date for each individual with Krabbe disease were simulated independently as the dates of randomly generated events in Poisson processes. Birth events were simulated by generating inter-arrival times between successive affected births independently drawn from an exponential distribution with rate parameter equal to *N* times *LR* until the diagnosis window was exceeded (note that the number of affected births simulated was always much less than *N*). Starting with the birthdate, each date of diagnosis was determined by drawing a corresponding age from an independent exponential distribution with rate parameter equal to $1 \div A$. In order to ensure that every single affected individual that could have survived into the diagnosis period was considered, birth and diagnosis dates were generated for the 110 y prior to the start of the diagnosis period.

Table 3. Parameter settings used during the Monte Carlo (MC) simulation

Parameter	Settings
Length of Diagnosis Period	5 – 30 y in increments of 5
Total Births Per Year	50,000 – 1,000,000 births in increments of 50,000
True Lifetime Risk at Birth	0.25, 0.5, 0.75, 1, 2, and 3 per 100,000
Expected Age at Diagnosis	0.3, 1, 2, 3, 4, 5

A series of MC simulation studies were conducted in which the total number of births per year, along with the birth dates and diagnosis dates of individuals with Krabbe disease, were simulated. Each MC simulation study was conducted using a distinct combination of parameter settings; the range of values used for each parameter is listed here. The range for each parameter was chosen to reflect the range of values observed in published studies of Krabbe disease.

For each affected birth, the probability of censoring was determined based on all-cause mortality rates in the US in 2009.[44] See **Supplemental Methods** for a detailed description of censoring calculations.

Lifetime risk calculations and criteria for comparison. Lifetime risk was estimated using the Dx, DOB and LT methods. The numerator of both the Dx method and the DOB method used the number of uncensored diagnoses observed in the diagnosis period. The Dx method divided this number by the total number of births (affected and unaffected) occurring within the diagnosis period (i.e., N × T). The DOB method, on the other hand, first required ascertainment of the earliest and latest birthdates of the diagnosed cases. The time between the earliest and latest birthdates is hereafter referred to as the "birth period." The number of total births occurring within the birth period (i.e., N × length of birth period) was divided into the number of observed diagnoses. The LT method is more intricate than the Dx and DOB methods and is described in detail elsewhere.[32] In addition to the information required by the Dx and DOB methods, the LT method requires the number of total deaths (due to all causes) per year and age at death for each member of the entire study population. The LT method uses this all-cause mortality data to estimate the number of censored cases.

Percent bias was calculated for each of the three estimation methods. Percent bias gives the percent over- or underestimation of a method as a percent of the true parameter value. Regression analysis was used to assess the associations between censoring, parameter settings and bias. The average percent bias of each method across all parameter settings was used to draw conclusions about the accuracy of each method.

MC results. The average percent bias scores for each method are listed in **Table 4**, and the results of the regression analyses can be found in **Table 5**. On average, the DOB method

Table 4. Results of the Monte Carlo (MC) Study: Accuracy of three estimation methods

% Bias	DOB	Dx	LT
1st Quartile	−42.50	−0.89	−0.49
Median	−24.60	−0.50	−0.12
Mean	−24.90	−0.51	−0.13
3rd Quartile	−7.73	−0.12	0.24

Percent bias describes the percent over- or underestimation of the true value, in this case the true lifetime risk at birth of Krabbe disease. Since the accuracy of an estimation method is influenced by factors that vary from study to study, many studies were simulated. How well each estimation method performed across all 2,160,000 simulated studies is described here. On average, the DOB method was significantly more inaccurate than either the Dx or LT method; and the Dx method was significantly more inaccurate than the LT method ($p < 0.01$ for all three comparisons, using paired t-tests and the Sidak correction for multiple comparisons).

underestimates true lifetime risk at birth by 24.9%, although this fluctuates dramatically based on the parameter settings used (significant associations between percent bias and each parameter setting, all $p < 0.001$). The Dx method underestimates by an average of 0.5%, with a significant association between censoring and percent bias ($p < 0.001$). The LT method is the most accurate, with an average underestimation of 0.1%, with a significant association between censoring and percent bias ($p < 0.05$) and between expected age at diagnosis and percent bias ($p < 0.001$).

Discussion and Recommendations for Future Research

Epidemiological measures are an important source of information for monitoring the morbidity and mortality of diseases in a population. They are also critical outcome measures for public health interventions. See the report of Macarov et al.[27] for an example in which epidemiological measures are used to document the success of a screening program for Krabbe disease implemented in a community in Israel. In the current review, we describe various problems that limit the utility of many other previously published epidemiological estimates of Krabbe disease rates in larger populations.

As described above, extant studies suffer from inappropriate terminology applied to epidemiological measures of Krabbe disease; in the current review, we show that nearly all previous estimates can be correctly interpreted as lifetime risk at birth. Previous studies often employ flawed methods to estimate lifetime risk at birth. In the current review, the DOB method was shown to be substantially inaccurate, while the Dx method and the LT method suffered from minimal error that was primarily attributable to censoring. The LT method was more accurate than the Dx method, as it is able to correct for most of the error introduced by censored cases. However, the difference in percent bias between the Dx and LT methods was small (less than one percentage point).

Due to the methodological issues discussed above, and given the Monte Carlo study results above, we recommend use of the LT method, when possible, to estimate the lifetime risk of Krabbe disease. In many cases, however, since the data required by the LT method is difficult to obtain in studies of Krabbe disease, the Dx method is an acceptable substitute as long as the population studied has a constant birth rate and diagnosis rate. The DOB method is inaccurate and should not be used. Aside from being the most accurate method, the LT method is superior for at least four other reasons: (1) it can yield estimates of remaining lifetime risk at ages other than birth; (2) it can be used to calculate estimates of standard error; (3) unlike the Dx method, the LT method is accurate whether or not the number of births per year remains constant; and (4) it is less affected by censoring. Since the association between censoring and bias is greater when the disease of interest arises later in life, the LT method is the only reasonable choice if the disease has a later onset.

As mentioned above, the between-study differences in case ascertainment, statistical methodology, and diagnostic capabilities are large enough that comparison of rates across studies is highly questionable. Future studies should pay special attention to case ascertainment, especially for older individuals with Krabbe disease who are more likely to be missed. If possible, unborn cases that have been diagnosed prenatally should be included in the estimated rate.

Age-specific mortality rates have been published for the United States,[38] but have not been published for any other country. Although not treated at length in the current paper, mortality rates provide useful information and should be similarly estimated and published for countries outside of the United States.

The results and conclusions of the current study are directly applicable to many disorders besides Krabbe disease. The DOB and Dx methods reviewed above have been applied in previous studies to various leukodystrophies,[17] sphingolipidoses,[21] lysosomal storage disorders,[7,18-20] and inborn errors of metabolism,[22,23] and all of the recommendations in the current paper are relevant to these disease categories as well.

Disclosure of Potential Conflicts of Interest

No potential conflict of interest was disclosed.

Supplemental Materials

Supplemental materials can be found here:
www.landesbioscience.com/journals/rarediseases/article/25212

Table 5. Regression results for bias analysis

	DOB		Dx		LT	
	Estimate	**p**	**Estimate**	**p**	**Estimate**	**p**
Intercept	−1.06e+02***	< 2e−16	−3.08e−02	0.59	−1.11e−01	0.05
Study/Birth Ratio	9.77e+01***	< 2e−16	-	-	–	-
Study Period Length	-	-	−6.89e−04	0.60	−6.59e−04	0.62
Total Births per Year	4.85e−03***	0.00058	−4.20e−04	0.29	−3.89e−04	0.33
True Lifetime Risk	2.66e−01***	6.2e−11	2.06e−02	0.069	2.09e−02	0.067
Expected Age at Diagnosis	7.22e−01***	< 2e−16	3.76e−03	0.79	5.93e−02***	0.000040
Percent Censored	2.02e+03***	< 2e−16	−4.73e+02***	5.2e−12	−1.53e+02*	0.026
R^2	99.03%		5.17%		0.79%	
	* p < 0.05	** p < 0.01	***p < 0.001			

Percent bias scores were analyzed separately for each estimation method (DOB, Dx, LT).

References

1. Hers HG. Inborn lysosomal diseases. Gastroenterology 1965; 48:625-33; PMID:14280390
2. Neufeld EF, Lim TW, Shapiro LJ. Inherited disorders of lysosomal metabolism. Annu Rev Biochem 1975; 44:357-76; PMID:806251; http://dx.doi.org/10.1146/annurev.bi.44.070175.002041
3. Luzi P, Rafi MA, Wenger DA. Structure and organization of the human galactocerebrosidase (GALC) gene. Genomics 1995; 26:407-9; PMID:7601472; http://dx.doi.org/10.1016/0888-7543(95)80230-J
4. Wenger DA. Krabbe Disease (Globoid Cell Leukodystrophy). In: Raymond GV, Eichler F, Fatemi A, Naidu S, eds. Leukodystrophies. London: Mac Keith Press; 2011.
5. Duffner PK, Caggana M, Orsini JJ, Wenger DA, Patterson MC, Crosley CJ, et al. Newborn screening for Krabbe disease: the New York State model. Pediatr Neurol 2009b; 40:245-52, discussion 253-5; PMID:19302934; http://dx.doi.org/10.1016/j.pediatrneurol.2008.11.010
6. Duffner PK, Jalal K, Carter RL. The Hunter's Hope Krabbe family database. Pediatr Neurol 2009a; 40:13-8; PMID:19068248; http://dx.doi.org/10.1016/j.pediatrneurol.2008.08.011
7. Poupetová H, Ledvinová J, Berná L, Dvoráková L, Kozich V, Elleder M. The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. J Inherit Metab Dis 2010; 33:387-96; PMID:20490927; http://dx.doi.org/10.1007/s10545-010-9093-7
8. Kolodny EH, Raghavan S, Krivit W. Late-onset Krabbe disease (globoid cell leukodystrophy): clinical and biochemical features of 15 cases. Dev Neurosci 1991; 13:232-9; PMID:1817026; http://dx.doi.org/10.1159/000112166
9. Jalal K, Carter R, Yan L, Barczykowski A, Duffner PK. Does galactocerebrosidase activity predict Krabbe phenotype? Pediatr Neurol 2012; 47:324-9; PMID:23044012; http://dx.doi.org/10.1016/j.pediatrneurol.2012.07.003
10. Adams RD, Lyon G. Practical laboratory tests for the diagnosis of hereditary metabolic encephalopathies. Neurology of Hereditary Metabolic Diseases of Children. Washington: McGraw-Hill; 1982.
11. Ozand PT, Gascon G, al Aqeel A, Roberts G, Dhalla M, Subramanyam SB. Prevalence of different types of lysosomal storage diseases in Saudi Arabia. J Inherit Metab Dis 1990; 13:849-61; PMID:2079833; http://dx.doi.org/10.1007/BF01800209
12. Lyon G, Hagberg B, Evrard P, Allaire C, Pavone L, Vanier M. Symptomatology of late onset Krabbe's leukodystrophy: the European experience. Dev Neurosci 1991; 13:240-4; PMID:1817027; http://dx.doi.org/10.1159/000112167
13. Arvidsson J, Hagberg B, Månsson JE, Svennerholm L. Late onset globoid cell leukodystrophy (Krabbe's disease)--Swedish case with 15 years of follow-up. Acta Paediatr 1995; 84:218-21; PMID:7756815; http://dx.doi.org/10.1111/j.1651-2227.1995.tb13616.x
14. Michelakakis H, Dimitriou E, Tsagaraki S, Giouroukos S, Schulpis K, Bartsocas CS. Lysosomal storage diseases in Greece. Genet Couns 1995; 6:43-7; PMID:7794561
15. Coelho JC, Wajner M, Burin MG, Vargas CR, Giugliani R. Selective screening of 10,000 high-risk Brazilian patients for the detection of inborn errors of metabolism. Eur J Pediatr 1997; 156:650-4; PMID:9266201; http://dx.doi.org/10.1007/s004310050685
16. Hagberg B, Kollberg H, Sourander P, Akesson HO. Infantile globoid cell leucodystrophy (Krabbe's disease). A clinical and genetic study of 32 Swedish cases 1953–1967. Neuropaediatrie 1969; 1:74-88; http://dx.doi.org/10.1055/s-0028-1091865
17. Heim P, Claussen M, Hoffmann B, Conzelmann E, Gärtner J, Harzer K, et al. Leukodystrophy incidence in Germany. Am J Med Genet 1997; 71:475-8; PMID:9286459; http://dx.doi.org/10.1002/(SICI)1096-8628(19970905)71:4<475::AID-AJMG20>3.0.CO;2-C
18. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. JAMA 1999; 281:249-54; PMID:9918480; http://dx.doi.org/10.1001/jama.281.3.249
19. Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, et al. Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 2004; 12:87-92; PMID:14685153; http://dx.doi.org/10.1038/sj.ejhg.5201044
20. Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, van Weely S, et al. The frequency of lysosomal storage diseases in The Netherlands. Hum Genet 1999; 105:151-6; PMID:10480370
21. Ozkara HA, Topçu M. Sphingolipidoses in Turkey. Brain Dev 2004; 26:363-6; PMID:15275696; http://dx.doi.org/10.1016/j.braindev.2003.09.006
22. Applegarth DA, Toone JR, Lowry RB. Incidence of inborn errors of metabolism in British Columbia, 1969-1996. Pediatrics 2000; 105:e10; PMID:10617747; http://dx.doi.org/10.1542/peds.105.1.e10
23. Dionisi-Vici C, Rizzo C, Burlina AB, Caruso U, Sabetta G, Uziel G, et al. Inborn errors of metabolism in the Italian pediatric population: a national retrospective survey. J Pediatr 2002; 140:321-7; PMID:11953730; http://dx.doi.org/10.1067/mpd.2002.122394
24. Bonkowsky JL, Nelson C, Kingston JL, Filloux FM, Mundorff MB, Srivastava R. The burden of inherited leukodystrophies in children. Neurology 2010; 75:718-25; PMID:20660364; http://dx.doi.org/10.1212/WNL.0b013e3181eee46b
25. Zlotogora J, Regev R, Zeigler M, Iancu TC, Bach G. Krabbe disease: increased incidence in a highly inbred community. Am J Med Genet 1985; 21:765-70; PMID:4025402; http://dx.doi.org/10.1002/ajmg.1320210420
26. Zlotogora J, Levy-Lahad E, Legum C, Iancu TC, Zeigler M, Bach G. Krabbe disease in Israel. Isr J Med Sci 1991; 27:196-8; PMID:2010272
27. Macarov M, Zlotogora J, Meiner V, Khatib Z, Sury V, Mengistu G, et al. Genetic screening for Krabbe disease: learning from the past and looking to the future. Am J Med Genet A 2011; 155A:574-6; PMID:21344630; http://dx.doi.org/10.1002/ajmg.a.33815
28. Suzuki K, Suzuki Y. Galactosylceramide Lipidosis: Globoid Cell Leukodystrophy (Krabbe's Disease). In: Stanbury JB, Wyngaarden JB, Fredrickson DS, Goldstein JL, Brown MS, eds. The Metabolic Basis of Inherited Disease, 5th edition. New York: McGraw-Hill; 1983.
29. Suzuki K, Suzuki Y, Suzuki K. Galactosylceramide lipidosis: globoid-cell leukodystrophy (Krabbe disease). In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Metabolic and molecular bases of inherited disease, 7th edition. New York: McGraw-Hill; 1995.
30. Wenger DA, Suzuki K, Suzuki Y, Suzuki K. Galactosylceramide lipidosis: globoid-cell leukodystrophy (Krabbe disease). In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease, 8th edition. New York: McGraw-Hill; 2001.
31. Pastores GM, Kolodny EH. Lysosomal storage diseases. In: Swaiman KF, Ashwal S, Ferriero DM, eds. Pediatric Neurology: Principles and Practice, 4th edition, Vol 1. Philadelphia: Elsevier; 2006.
32. Beiser A, D'Agostino RB Sr., Seshadri S, Sullivan LM, Wolf PA. Computing estimates of incidence, including lifetime risk: Alzheimer's disease in the Framingham Study. The Practical Incidence Estimators (PIE) macro. Stat Med 2000; 19:1495-522; PMID:10844714; http://dx.doi.org/10.1002/(SICI)1097-0258(20000615/30)19:11/12<1495::AID-SIM441>3.0.CO;2-E

33. Rothman KJ. Modern Epidemiology. Boston: Little, Brown and Company; 1986.
34. MacMahon B, Trichopoulos D. Epidemiology: Principles and Methods, 2nd edition. Boston: Little, Brown and Company; 1996.
35. Rothman KJ. Epidemiology: An Introduction. Oxford: University Press; 2002.
36. Koepsell TD, Weiss NS. Epidemiologic methods: Studying the outcome of illness. Oxford: University Press; 2003.
37. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. Lancet Neurol 2007; 6:1106-14; PMID:18031707; http://dx.doi.org/10.1016/S1474-4422(07)70291-0
38. Barczykowski AL, Foss AH, Duffner PK, Yan L, Carter RL. Death rates in the U.S. due to Krabbe disease and related leukodystrophy and lysosomal storage diseases. Am J Med Genet A 2012; 158A:2835-42; PMID:22991292; http://dx.doi.org/10.1002/ajmg.a.35624
39. Metropolis N, Ulam S. The Monte Carlo method. J Am Stat Assoc 1949; 44:335-41; PMID:18139350; http://dx.doi.org/10.1080/01621459.1949.10483310
40. de Vladar HP, Barton NH. The contribution of statistical physics to evolutionary biology. Trends Ecol Evol 2011; 26:424-32; PMID:21571390; http://dx.doi.org/10.1016/j.tree.2011.04.002
41. Pawlicki T, Ma CM. Monte Carlo simulation for MLC-based intensity-modulated radiotherapy. Med Dosim 2001; 26:157-68; PMID:11444518; http://dx.doi.org/10.1016/S0958-3947(01)00061-9
42. Broadie M, Glasserman P. Estimating security price derivatives using simulation. Manage Sci 1996; 42:269-86; http://dx.doi.org/10.1287/mnsc.42.2.269
43. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2011.
44. Martin JA, Hamilton BE, Ventura SJ, Osterman MJ, Kirmeyer S, Mathews TJ, et al. Deaths: final data for 2009. Natl Vital Stat Rep 2011; 60:1-116.
45. Claussen M, Heim P, Knispel J, Goebel HH, Kohlschütter A. Incidence of neuronal ceroid-lipofuscinoses in West Germany: variation of a method for studying autosomal recessive disorders. Am J Med Genet 1992; 42:536-8; PMID:1609834; http://dx.doi.org/10.1002/ajmg.1320420422
46. Houston J, Aschlimann M, Barrett M, Robison C. Births and Deaths in 2009. Utah Office of Vital Records and Statistics. Technical Report No. 271. 2011.