

Individual heterogeneity as a pitfall in population estimates based on non-invasive genetic sampling: a review and recommendations

Cornelia Ebert, Felix Knauer, Ilse Storch & Ulf Hohmann

In recent years, much progress has been made in non-invasive genetic methods for various purposes including population estimation. Previous research focused on optimising laboratory protocols and assessing genotyping errors. However, an important source of bias in population estimates still remains in the field sampling methods. The probability of animals being sampled can vary according to sex, age, social status or home-range location. In this article, we present relevant literature reviewed to provide an overview of the occurrence of individual heterogeneity (IH) in the field, and how it can be minimised, e.g. by adaptation of sampling design. We surveyed 38 articles describing non-invasive population estimation for 12 mammal and two bird species. The majority of these studies discussed IH as a potential problem. The detectability of IH via goodness-of-fit testing depended on the average capture probability reported in the studies. Field tests for assessing variation in sampling probabilities or validating estimations were carried out in only 11 of the 38 studies. The results of these tests point out that IH is a widespread problem in non-invasive population estimation, which deserves closer attention not only in the development of laboratory protocols but also concerning the sampled species' characteristics and the field methods. IH can be reduced in the field by carefully adapting the sampling design to the characteristics of the studied population. If this is not reasonable, it may be better to switch to a different sampling strategy.

Key words: *Canis latrans, Canis lupus, capture-recapture, faeces, genotyping, hair sampling, individual heterogeneity, Meles meles, population estimate, Rhinolophus hipposideros, Sus scrofa, Ursus americanus, Ursus arctos, wildlife management*

Cornelia Ebert & Ulf Hohmann, Department of Forest Ecology and Wildlife, Research Institute for Forestry and Forest Ecology of Rhineland-Palatinate, Hauptstrasse 16, D-67705 Trippstadt, Germany - e-mail addresses: cebert@gmx.de (Cornelia Ebert); ulf.hohmann@wald.rlp.de (Ulf Hohmann)

Felix Knauer, Department of Wildlife Ecology and Management, University of Freiburg, Tennenbacher Strasse 4, D-79106, Freiburg, Germany - e-mail: felix.knauer@fiwi.at*

Ilse Storch, Department of Wildlife Ecology and Management, University of Freiburg, Tennenbacher Strasse 4, D-79106 Freiburg, Germany - e-mail: ilse.storch@wildlife.uni-freiburg.de

**Present address: Research Institute for Wildlife Ecology, Savoyenstrasse 1, A-1160 Vienna, Austria*

Corresponding author: Cornelia Ebert

Received 9 December 2009, accepted 28 June 2010

Associate Editor: Olivier Gimenez

Reliable estimation of population size remains a major challenge in wildlife research and management. In recent years, non-invasive DNA-based population estimation methods have been widely applied in a variety of species. Several standard approaches have been modified to fit genetical implementation, among which are rarefaction (e.g. Frantz et al. 2004) and capture-recapture (e.g.

Boulanger & McLellan 2001). In their conventional form, both methods presuppose capture or killing of animals or rely on direct sightings, and are challenged by the possibility of heterogeneous detection probabilities amongst the studied population (Borchers et al. 2002, Petit & Valière 2006). Being most frequently used, capture-recapture (CR) methods are especially vulnerable with respect to

individual heterogeneity (IH; Pledger & Efford 1998, Link 2004, Lukacs & Burnham 2005), i.e. differences between individuals of a population in the probability of being captured (Borchers et al. 2002). Capture- and recapture probabilities may be influenced by age, sex, social status and individual experience (Baber & Coblenz 1986, Piggott & Taylor 2003), and this can generate severe bias in population estimates (White et al. 1982, Minta & Mangel 1989, Sweitzer et al. 2000). IH can be accounted for by use of different modelling approaches (e.g. see Otis et al. 1978, Chao 1987, Chao & Jeng 1992, Pledger & Efford 1998), but the power of goodness-of-fit (GOF) tests and model selection procedures to detect IH in a given data set is often low (Menkens & Anderson 1988, McKelvey & Pearson 2001). Furthermore, as Link (2003, 2004) recently stated, IH is far more difficult to model than has previously been recognised, modelling being especially problematic if the causes and extent of IH are unknown. Thus, in order to allow accurate population estimates, IH should be either minimised or quantified as far as possible (Petit & Valière 2006).

Methods based on non-invasive genetic sampling offer solutions for estimation of population size without capturing or killing animals, making them advantageous for rare or endangered species (Kohn et al. 1999, Taberlet et al. 1999, Mills et al. 2000, Piggott & Taylor 2003). McKelvey & Schwartz (2004) and Petit & Valière (2006) suggested that the absence of handling can overcome the effects of previous capture history on subsequent catchability, and thus certain sources of IH could be reduced. The most commonly used non-invasive DNA sources are hairs and faeces for mammals as well as feathers and faeces for birds (Lukacs & Burnham 2005). Non-invasive methods have made CR approaches, which in their conventional form are more suitable for small and abundant mammals, applicable for large, elusive and/or endangered mammal and bird species (Obbard et al. 2010).

However, despite their advantages, non-invasive genetic methods are also prone to heterogeneity related to biological variability among individuals (Kohn et al. 1999, Wilson et al. 2003, Boulanger et al. 2004a). Moreover, in non-invasive methods IH can interact with bias caused by genotyping errors. Allelic dropout and false alleles can create 'new' false individuals, leading to overestimation in population estimates because recaptures may be concealed, resulting in a decreased recapture rate

(Creel et al. 2003, McKelvey & Schwartz 2004). Furthermore, there are some issues in non-invasive genetic CR which are not problematic in conventional CR. In genetic CR the total number of marks in the population is not known and marks may not be unique, because only a subset of each animal's genome is used for identification (Lukacs & Burnham 2005). Therefore, the danger of misidentification is increased compared to conventional CR. Also, a 'sampling occasion' can be more difficult to define than a 'capture occasion', because the moment of the deposition of a sample, e.g. hair or faeces, can not be assessed precisely. This can compromise the concept of population closure (Lukacs & Burnham 2005). Thus, despite the high potential of non-invasive genetic techniques, there are several issues which can complicate the application of a CR framework for population estimation, in addition to the difficulties already present in the conventional approach.

Until now, great progress has been made in genetic techniques. In particular, much effort has been devoted to quantifying and reducing genotyping errors (Taberlet et al. 1999, Paetkau 2003, Broquet & Petit 2004, Roon et al. 2005b, Miquel et al. 2006). In contrast, fewer attempts have been made in assessing the extent and causes of IH in the field, i.e. due to biological characteristics of the sampled species, to individual attributes or due to sampling procedures (Boulanger et al. 2006). However, information about the causes and the extent of IH is essential to improve sampling designs (Boulanger et al. 2004c). Furthermore, IH in combination with uncertainties caused by genotyping errors can cause multiplicative effects and thus lead to an increase in the overall bias. Therefore, it is crucial to address both IH and genotyping problems very carefully in order to minimise bias in population estimates.

Based on the recent peer-reviewed literature, our article aims at:

- 1) providing a survey of the occurrence and treatment of IH in non-invasive population estimation studies, especially with respect to different sampling strategies;
- 2) assessing the impact of sample size and capture probability (p) on the detectability of IH via GOF tests or model selection procedures in CR studies;
- 3) comparing different methods, which seem suitable to assess IH in the field, also with respect to the study species and its characteristics.

Material and methods

Our review is based on population genetic studies that involve non-invasive sampling for the purpose of population estimation in mammal and bird species. We performed a search in the Swiss Wildlife Information Service (SWIS) database for peer-reviewed publications using the following search terms: hair trap, non-invasive sampling, genotyping, population estimates, faeces sampling, hair sampling and genetic monitoring. The search yielded 104 titles and was supplemented with published lists of references. In total, we detected 142 articles, of which we focused on 38 studies (complete list of references available on request from the corresponding author). We only included papers in which the non-invasive sampling was *de facto* conducted in the field and applied for population estimation; we excluded literature reviews and articles dealing with single aspects in the development of sampling methods. We focused on studies using hair, faeces and feathers as those are the main sources of non-invasive tissue samples. Other sources (e.g. urine, shed skin or buccal swabs) have been much less employed for population estimation until now (Broquet et al. 2007). We also included cases in which a combination of different sampling strategies was applied. We restricted our review to studies using CR or rarefaction (also termed 'accumulation curve' methods; e.g. see Kohn et al. 1999 and Eggert et al. 2003) approaches for population estimation, as they are the most commonly used and more prone to be biased by IH than e.g. estimation of minimum densities or minimum number alive.

For each study, we assessed whether IH had been mentioned, i.e. considered as a factor potentially influencing the population estimate. Additionally, we recorded if IH was detected, e.g. via likelihood-ratio-tests (program CAPWIRE; Miller et al. 2005), χ^2 -tests (program CAPTURE; Otis et al. 1978), via Akaike's Information Criterion (AIC, e.g. in program MARK; White & Burnham 1999) or GOF testing in program U-CARE (Choquet et al. 2009). Furthermore, IH can be discerned in uneven 'capture frequencies' of sampled individuals (Kohn et al. 1999, Scheppers et al. 2007). The power of tests to detect IH can depend on capture probability (p ; Pollock et al. 1990, Boulanger et al. 2002). Additionally, we suspected the number of sampling occasions and coverage, i.e. proportion of the population sampled, to have an effect on IH detectability. The estimated coverage is significantly

correlated with p and was included because not all reviewed studies provided estimates of p . We used logistic regression to evaluate the impact of p , coverage and sampling occasions on the probability of IH being detected. In this context, we evaluated studies in which IH was detected in the capture frequency or via field test, but not in the GOF or model selection tests as 'not detected'. We also included squared terms of p and coverage since data suggested an optimum somewhere in between the extreme values. We selected models based on AIC. For the logistic regression, we used every single population estimate reported in the 38 reviewed articles, as results of several study years or different study areas were often included in one article, resulting in different p and population estimates. As p and coverage of different analyses reported within the same study could be correlated, we included the studies as random factor. This worked well with coverage but not with p , because for p sample size was low (only 39 of the 76 analyses included estimates of p average) and the number of studies reporting only one analysis was high. Therefore, in the case of p we averaged the values for each study and conducted both a weighted and an unweighted logistic regression without random effects. All analyses were performed using program R (Ihaka & Gentleman 1996). For the mixed effect logistic regression model we used the function `lmer` of the package `lme4` (Bates & Maechler 2010).

For studies using CR methodology, we additionally recorded if IH was included in the estimation model (which is not possible for rarefaction methods). We also searched for studies in which field tests had been carried out parallel to the non-invasive sampling for validation purposes. We put our special attention to methods and results of these studies and aimed to assess if the applied field tests hold the potential to reveal IH.

Results

The articles we reviewed dealt with 14 different study species; 12 mammals and two birds. In 30 of the 38 studies we included in our review, CR was the sole method applied to estimate population size. Four studies used rarefaction analysis only and four used both methods (Table 1). Altogether, hair was the DNA source in 22 of the 38 studies. In one of these cases, the hair sampling was combined with harvest data, and in another one faecal sampling

Table 1. Details about all the reviewed articles. For each article, every single population estimate is registered separately (e.g. in studies which were carried out over several years or in different study areas). P average is the average per session capture probability reported in the studies (n.r. = not reported). For each study, it is noted if individual heterogeneity (IH) was detected via goodness-of-fit (GOF) testing or model selection procedure.

Reference (study)	Species	Sampling	No. of individuals sampled	P average	Estimated N	No. of sampling occasions	Software used	IH detected in the field? (test method)	IH detected via GOF (test)
Banks et al. 2003	<i>Lastorhinus kreffii</i>	hair	81	n.r.	113	7	CAPTURE	-	yes, CF (none)
Mowat & Paetkau 2002	<i>Martes americana</i>	hair	88	0.15	213	4	CAPTURE	-	yes (χ^2)
Scheepers et al. 2007	<i>Meles meles</i>	hair	55	n.r.	61	1	CAPWIRE	yes (direct sightings)	yes (LR)
Triant et al. 2004 (coast sampling)	<i>Ursus americanus</i>	hair	57	n.r.	77	2	n.r.	-	no (-)
Triant et al. 2004 (inland sampling)	<i>Ursus americanus</i>	hair	32	n.r.	41	2	n.r.	-	no (-)
Boersen et al. 2003	<i>Ursus americanus</i>	hair	58	n.r.	119	14	CAPTURE	-	yes (χ^2)
Immell & Anthony 2006 (Steamboat 2003)	<i>Ursus americanus</i>	hair	32	0.4	46	3	CAPTURE	-	yes (χ^2)
Immell & Anthony 2006 (Steamboat 2004)	<i>Ursus americanus</i>	hair	30	0.2	57	3	CAPTURE	-	no (χ^2)
Immell & Anthony 2006 (Tokete 2003)	<i>Ursus americanus</i>	hair	47	0.3	67	3	CAPTURE	-	yes (χ^2)
Immell & Anthony 2006 (Tokete 2004)	<i>Ursus americanus</i>	hair	46	0.31	65	3	CAPTURE	-	yes (χ^2)
Belant et al. 2005 (Stockton Island)	<i>Ursus americanus</i>	hair	26	0.68	26	4	CAPTURE	-	no (χ^2)
Belant et al. 2005 (Sand Island)	<i>Ursus americanus</i>	hair	6	0.54	6	4	CAPTURE	-	sample size too small for test
Settlage et al. 2008 (National Park)	<i>Ursus americanus</i>	faeces	129	0.06 - 0.18	97 - 114	10	CAPTURE	-	yes (χ^2)
Settlage et al. 2008 (National Forest)	<i>Ursus arctos</i>	faeces	60	0.09 - 0.32	197 - 330	10	CAPTURE	-	yes (χ^2)
Poole et al. 2001 ¹	<i>Ursus arctos</i>	hair	98	0.19	148	5	CAPTURE	-	no (χ^2)
Boulanger & McLellan 2001 ¹ (CAPTURE) ²	<i>Ursus arctos</i>	hair	98	n.r.	155	5	CAPTURE	-	no (χ^2)
Boulanger & McLellan 2001 (MARK) ¹	<i>Ursus arctos</i>	hair	98	0.05 - 0.4	191	5	MARK	-	yes (AIC)
Boulanger et al. 2002 (Jumbo project)	<i>Ursus arctos</i>	hair	33	0.26	45	4	CAPTURE	-	yes (χ^2)
Boulanger et al. 2002 (U. Columbia 97)	<i>Ursus arctos</i>	hair	40	0.2	55	5	CAPTURE	-	yes (χ^2)
Boulanger et al. 2002 (U. Columbia 98)	<i>Ursus arctos</i>	hair	40	0.12	92	5	CAPTURE	-	yes (χ^2)
Boulanger et al. 2002 (Kingcome)	<i>Ursus arctos</i>	hair	58	0.2	102	5	CAPTURE	-	yes (χ^2)
Boulanger et al. 2002 (U. Columbia 96)	<i>Ursus arctos</i>	hair	55	0.16	108	4	CAPTURE	-	yes (χ^2)
Boulanger et al. 2002 (Granby Kettle)	<i>Ursus arctos</i>	hair	22	0.13	46	5	CAPTURE	-	no (χ^2)
Boulanger et al. 2002 (Prophet)	<i>Ursus arctos</i>	hair	98	0.17	166	5	CAPTURE	-	yes (χ^2)
Boulanger et al. 2004b	<i>Ursus arctos</i>	hair	41	0.35	104	3	MARK	yes (radio-telemetry)	yes (AIC)

Table 1. Continued.

Reference (study)	Species	Sampling	No. of individuals sampled	P average	Estimated N	No. of sampling occasions	Software used	IH detected in the field? (test method)	IH detected via GOF (test)
Boulanger et al. 2004a (sampling 1996) (radiotelemetry)	<i>Ursus arctos</i> no (AIC)	hair	54	0.16	108	4	CAPTURE	yes	
Boulanger et al. 2004a (sampling 1997)		hair	41	0.2	55	5	CAPTURE	no (AIC)	no (AIC)
Boulanger et al. 2004a (sampling 1998)		hair	39	0.12	92	5	CAPTURE	no (AIC)	
Boulanger et al. 2006	<i>Ursus arctos</i>	hair	41	0.32	43	4	MARK	-	yes (AIC)
Boulanger et al. 2004c	<i>Ursus arctos</i>	hair	total: 123	n.a.	per year, mean 49	5	MARK	-	no (-)
Mowat & Strobeck 2000 (British Columbia)	<i>Ursus arctos</i>	hair	109	0.1	257	5	CAPTURE	-	no (χ^2)
Mowat & Strobeck 2000 (Alberta)		hair	37	0.16	74	4	CAPTURE	-	yes (χ^2)
Mowat et al. 2005 (sampling SC Selkirks)	<i>Ursus arctos</i>	hair	38	0.09	97	5	CAPTURE	-	?
Mowat et al. 2005 (sampling NC Selkirks)		hair	74	0.08	223	5	CAPTURE	-	?
Mowat et al. 2005 (sampling Prophet plateau)		hair	32	0.13	63	5	CAPTURE	-	no (χ^2)
Mowat et al. 2005 (sampling Prophet Mtns)		hair	67	0.21	96	5	CAPTURE	-	no (χ^2)
Mowat et al. 2005 (sampling Yellowhead)		hair	48	0.16	107	4	CAPTURE	-	?
Mowat et al. 2005 (sampling Parsnip Plateau)		hair	21	0.12	50	4	CAPTURE	-	no (χ^2)
Mowat et al. 2005 (sampling Parsnip Mtns)		hair	216	0.22	341	4	CAPTURE	-	no (χ^2)
Mowat et al. 2005 (sampling Parsnip Plateau)	<i>Ursus americanus</i>	hair	194	0.06	892	4	CAPTURE	-	?
Mowat et al. 2005 (sampling Parsnip Mtns)		hair	85	0.08	363	4	CAPTURE	-	?
Mowat et al. 2005 (sampling Bowron river)	<i>Ursus arctos</i>	hair	53	0.32	76	3	CAPTURE	-	?
Woods et al. 1999	<i>Ursus arctos</i> , <i>U.</i>	hair	54	5	104	4	CAPTURE	-	no (-)
Gervasi et al. 2008 ³	<i>americanus</i> <i>Ursus arctos</i> <i>mariscanus</i>	hair	30	0.03	44	12	MARK	-	no (-)
Frantz et al. 2004	<i>Meles meles</i>	hair	14	n.r.	12 - 19	n.r.	GIMLET/ R (rarefaction)	no (direct counts)	no (-)
Dreher et al. 2007	<i>Ursus americanus</i>	hair + harvested bears	544	0.02 (hair), 0.21 (hunt)	1882	5	MARK	radio-telemetry, harvested bears	no (-)
Wasser et al. 2004	<i>Ursus arctos</i> , <i>U.</i> <i>americanus</i>	hair, faeces	24	n.r.	28	5	MARK	(hair: yes) radio-telemetry, comparison between sampling methods (no)	no (-)

Table 1. Continued.

Reference (study)	Species	Sampling	No. of individuals sampled	P average	Estimated N	No. of sampling occasions	Software used	IH detected in the field? (test method)	IH detected via GOF (test)
Cubaynes et al. 2010	<i>Canis lupus</i>	faeces	160	0.01 - 0.86	3 - 126	n.r.	E-SURGE, U-CARE	-	yes (AIC)
Prugh et al. 2005	<i>Canis latrans</i>	faeces	total: 56	0.6 and 0.75	per year, mean 26	9	MARK	radio-telemetry (yes)	yes (AIC, CF)
Prigioni et al. 2006	<i>Lutra lutra</i>	faeces	23	n.r.	36	n.r.	SPSS (rarefaction)	-	yes, CF (none)
Puechmaille & Petit 2007 (sampling Epiniac 2003)	<i>Rhinolophus hipposideros</i>	faeces	54	n.r.	approx. 65	1	R	direct counts (no)	no (LR, S)
Puechmaille & Petit 2007 (sampling Pluherlin 2003)		faeces	35	n.r.	approx. 42	1	R		no (LR, S)
Puechmaille & Petit 2007 (sampling Saint-Thurial 2003)		faeces	16	n.r.	approx. 28	1	R		yes (LR, S)
Puechmaille & Petit 2007 (sampling Epiniac 2004)		faeces	58	n.r.	approx. 85	1	R		yes (LR, S)
Puechmaille & Petit 2007 (sampling Pluherlin 2004)		faeces	35	n.r.	approx. 48	1	R		yes (LR, S)
Puechmaille & Petit 2007 (sampling Saint-Thurial 2004)		faeces	14	n.r.	approx. 18	1	R		no (LR, S)
Frantz et al. 2003 ⁴	<i>Meles meles</i>	faeces	20	0.15	26	10	CAPTURE	-	n.r.
Bellemain et al. 2005 (sampling 2001)	<i>Ursus arctos</i>	faeces	311	n.r.	approx. 480	11	MARK	-	yes (AIC)
Bellemain et al. 2005 (sampling 2002)		faeces	239	n.r.	approx. 350	13	MARK	-	yes (AIC)
Bellemain et al. 2007	<i>Ursus arctos</i>	faeces	28	n.r.	32 and 47	n.r.	GIMLET/ R (rarefaction)	-	no (-)
Solberg et al. 2006 (sampling 2001)	<i>Ursus arctos</i>	faeces	146	n.r.	223	11	MARK	-	yes (AIC)
Solberg et al. 2006 (sampling 2002)		faeces	81	n.r.	157	13	MARK	-	yes (AIC)
Jacob et al. 2009 (sampling Obwalden)	<i>Tetrao urogallus</i>	faeces	29	n.r.	78	1	CAPWIRE	-	yes (LR)
Jacob et al. 2009 (sampling Regelstein)		faeces	9	n.r.	14	1	CAPWIRE	-	yes (LR)
Jacob et al. 2009 (sampling Höhi)		faeces	16	n.r.	20	1	CAPWIRE	-	no (LR)
Jacob et al. 2009 (sampling Schwägalp)	<i>Tetrao urogallus</i>	faeces	7	n.r.	10	1	CAPWIRE	-	yes (LR)
Jacob et al. 2009 (sampling Rofla)		faeces	7	n.r.	10	1	CAPWIRE	-	yes (LR)
Jacob et al. 2009 (sampling Salouf)		faeces	5	n.r.	5	1	CAPWIRE	-	no (LR)
Jacob et al. 2009 (sampling Albulanorth)		faeces	23	n.r.	36	1	CAPWIRE	-	no (LR)
Jacob et al. 2009 (sampling Albulan-south)		faeces	8	n.r.	33	1	CAPWIRE	-	no (LR)
Eggert et al. 2003	<i>Loxodonta cyclotis</i>	faeces	86	0 - 0.2	225	10	CAPTURE	-	yes (χ^2)
Banks et al. 2002	<i>Vombatus ursinus</i>	faeces	17	n.r.	19	5	CAPTURE	-	yes (χ^2)

Table 1. Continued.

Reference (study)	Species	Sampling	No. of individuals sampled	P average	Estimated N	No. of sampling occasions	Software used	IH detected in the field? (test method)	IH detected via GOF (test)
Rudnick et al. 2008	<i>Aquila heliaca</i>	faeces, feathers	278	> 0.44	308	4	MARK	-	yes, CF (none)
Marucco et al. 2009 ⁵	<i>Canis lupus</i>	faeces	total: 87	0.28 - 0.77	per year, 21 - 47	14	MSURGE, UCARE	yes ⁵ (snow tracking, marking behaviour)	no (AIC)
Meijer et al. 2008	<i>Alopex lagopus</i>	faeces	30	n.r.	42	2	n.r.	-	no (-)
Wilson et al. 2003 ⁴	<i>Meles meles</i>	faeces	20	n.r.	36	4	Proc NLIN	video control, marked badgers (yes)	yes, CF (none)
Kohn et al. 1999	<i>Canis latrans</i>	faeces	30	n.r.	41	n.r.	JMP IN3	radio-telemetry (yes)	yes, CF (none)

LR indicates Likelihood-ratio test (see Miller et al. 2005) implemented in model selection procedure of program CAPWIRE;

S indicates simulation test;

χ^2 indicates χ^2 -tests implemented in Program CAPTURE model selection routine (Otis et al. 1978);

AIC indicates IH suggested by the model selection process (via Akaike's Information Criterion) in program MARK (White & Burnham 1999);

CF indicates capture frequency, i.e. IH is visible in the number of times different individuals are sampled.

¹ Both articles deal with the same grizzly bear data set (Prophet River Project 1998). In both, χ^2 -tests failed to detect IH. But Boulanger & McLellan (2001) detected IH as a function of distance from study area (via AIC) as well as assumed age-specific capture probabilities.

² For Boulanger & McLellan (2001), two different population estimation methods are reported for the same data set. Because of their differing results, both are listed here separately.

³ In this paper, heterogeneity was modelled for each of three different hair sampling approaches applied, but not IH as discussed in our review due to small sample size. Thus, IH was assumed by the authors to be present in the data set, but not detected via tests.

⁴ Both articles deal with the same badger data set. In Frantz et al. (2003), results of program CAPTURE's χ^2 -tests are not reported, even though the authors note that they assume the heterogeneity models to be most appropriate. Wilson et al. (2003) reported the capture frequencies of individual badgers which revealed IH.

⁵ IH was detected, but it did not impact the population estimate; for details see text.

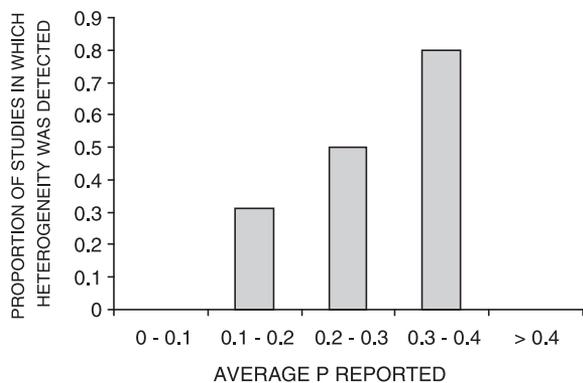


Figure 1. Proportion of population estimates in which individual sampling heterogeneity was detected via goodness-of-fit test in relation to average capture probability (average p reported). Only studies which report an estimate of the average capture probability p are included (39 of 75 studies; see Table 1).

was carried out simultaneously. The remaining 16 studies relied on faeces as DNA source; one of them in combination with feathers. IH bias was mentioned as a potential problem in 34 (89%) of the 38 contributions, whereas it was modelled in 26 of the 34 CR studies (i.e. 76%). In 18 of the 34 CR articles, IH was detected via χ^2 -tests, likelihood-ratio-tests or with the help of AIC. In another five studies in which GOF tests were performed, IH was not detected in their data sets. In six of all the 38 studies, tests either failed to detect IH or were not performed, but it was nevertheless visible in the 'capture frequencies' (see Table 1). Altogether, in 24 of the 38 studies (63%) the data revealed the occurrence of IH amongst the studied population independent of further field tests.

Assessment of IH via GOF tests and model selection procedures

In this section we are taking into account every single population estimate ($N = 76$) reported in the

Table 2. Support of logistic regression models testing the impact of coverage (i.e. the ratio on the detection of individual heterogeneity in the reviewed studies ('samp_occ' = number of sampling occasions).

Model	AIC		AIC	
	weighted	Δ AIC	unweighted	Δ AIC
p, p^2	27.49	0	14.30	0
p, p^2 , samp_occ	29.40	1.90	16.12	1.82
p, p^2 , samp_occ, p^* samp_occ, p^2 *samp_occ	32.64	5.15	19.80	5.50
Null	36.00	8.51	24.08	9.78
samp_occ	37.45	9.95	25.81	11.51
P	37.45	9.96	24.41	10.11

Table 3. Support of logistic regression models testing the impact of capture probability on the detection of individual heterogeneity in the reviewed studies ('p' = capture probability, ' p^2 ' = squared capture probability, 'samp_occ' = number of sampling occasions).

Model	AIC	Δ AIC
p, p^2	27.49	0
p, p^2 , samp_occ	29.40	1.90
p, p^2 , samp_occ, p^* samp_occ, p^2 *samp_occ	32.64	5.15
Null	36.00	8.51
samp_occ	37.45	9.95
P	37.45	9.96

38 reviewed articles. More than half (53.8%) of the 39 reported estimates of p average are below 0.2 (Fig. 1) and thus around or below the minimum recommended by Otis et al. (1978) for reliable model selection and population estimates (Otis et al. 1978 recommend $p \geq 0.2$ for a population of 200 animals and state that p should never be below 0.1. Recommended minimum number of capture occasions is 5, but better 7 - 10). Furthermore, IH was only detected for p between 0.16 and 0.4. The proportion of population estimates in which IH was detected increased with increasing p until $p = 0.4$ (see Fig. 1). In none of three studies with $p \geq 0.4$ IH was detected. However, in one of these studies the sample size was too small to carry out tests in the program CAPTURE (see Table 1; Belant et al. 2005 sampling on Sand Island). The logistic regression showed no impact of coverage on the detectability of IH (Table 2). In the case of p and p^2 , the results suggest that there is an effect on the detectability of IH (Tables 3 and 4). The IH detectability is highest at p values around 0.3 (Fig. 2). The most supported model does not include the number of sampling occasions, but a model including sampling occasions is ranked marginally below (Δ AIC < 2), indicating a potential influence (Burnham & Anderson 1998). It seems possible that with an increasing number of sampling occasions, the detectability of IH increases. Models including an interaction between p and the

Table 4. Results of the regression model testing for the impact of capture probability on the detectability of individual heterogeneity with best fit according to AIC ranking (model 'p, p^2 ').

Parameter	Estimate	SE	P
intercept	-12.071	5.856	0.0393
p	96.611	49.707	0.0519
p^2	-161.981	89.441	0.0701

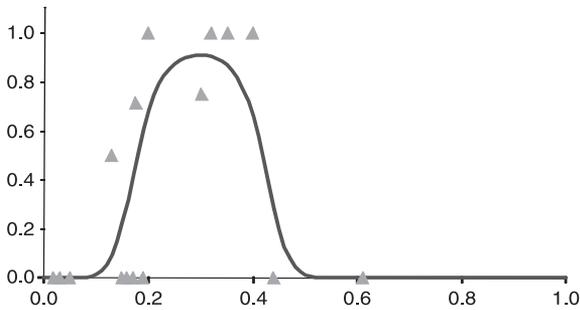


Figure 2. Observed (▲) and predicted (—) values of IH detectability dependent on p and p^2 .

number of sampling occasions were not supported (see Table 3).

Assessment of IH via field tests

Field tests suitable for assessing the occurrence of IH bias were performed in 11 of the 38 studies. In seven of these 11 cases, IH was actually detected (see Table 1). In one of the four other cases, IH was found to be present in the hair sampling part of the study, but was strongly reduced by sampling harvested animals as an additional strategy (Dreher et al. 2007). Furthermore, in eight of the 11 studies, IH was detected via GOF testing or in the 'capture frequencies'. Thus, in two cases where the field tests did not reveal IH, nevertheless it seemed to be present and detectable in the data set. Furthermore, in two cases IH was detected through the field test but not in the data.

Field tests in detail

Using radio-telemetry, Kohn et al. (1999) found IH to be present in their population under study: 12 radio-collared coyotes *Canis latrans* made use of the area to different degrees, and the number of faeces that was deposited correlated with their relative use of the study area. This IH was also reflected in the 'capture frequencies' of the sampled individuals. Faeces sampling of a coyote population in Central Alaska, USA, exhibited IH with respect to age and home range, as well as resident status. This was revealed through radio-telemetry of 15 collared adult resident individuals, which showed higher survival and recapture rates than juveniles and transient or edge individuals (Prugh et al. 2005). Furthermore, the model selection process in program MARK detected IH in the data. This was also the case in a hair sampling study on grizzly bears *Ursus arctos*; Boulanger et al. (2004c) used location data of 12 GPS-collared bears to evaluate the po-

tential bias. They found p to be greater for males than for females and also to be influenced by capture history (i.e. differences between collared and non-collared individuals). The latter was also detected in another study (Boulanger et al. 2004a) in which radio-telemetry was conducted over three years on a total of 35 bears and were compared to hair sampling data collected in the same area. Additionally, the p of females with cubs differed from those of the rest of the population. In a study carried out by Wasser et al. (2004), grizzly bear faeces sampling data were compared to hair sampling and radio-telemetry data collected simultaneously in the same area. Faeces collection was conducted with the help of trained dogs. This seemed to be an effective and relatively less biased method than hair sampling at baited stations. In the latter, close kin (i.e. females with their offspring) were considerably less represented. However, the sample size of matched faeces and telemetry data was too small to allow more fine-grained comparisons, and IH was neither detected in the field test nor in the data set (Wasser et al. 2004). In a black bear *Ursus americanus* hair sampling study accompanied by radio-telemetry, Dreher et al. (2007) used harvested bears as an additional sample. Due to this combination, the IH, which would have been present if hair sampling or harvest data were used alone, was strongly reduced (B. Dreher, pers. comm.) and thus neither detected in the field test nor in the data set.

Wilson et al. (2003) carried out a faeces sampling study on badgers *Meles meles* in which they used video control of a largely marked population to validate their rarefaction estimate. They succeeded in sampling almost the entire population by collecting faeces at latrines near badger setts and did not detect sex- or age class bias. However, considerable variation existed in the numbers of samples obtained from the different individuals. Moreover, some known individuals never used the sampled latrines and thus were not identified via faeces sampling. These results suggest the incidence of IH, e.g. due to variation in individual behaviour, which in that case might not have compromised the estimation because such a high proportion of the population was actually sampled. A hair sampling study conducted by Scheppers et al. (2007) at badger setts, simultaneously surveyed via direct observations, yielded similar results. The ease with which badgers were sampled varied considerably between setts; hair traps were not visited equally often by all members of the groups. Using baited hair traps and

applying direct observation as a validation method, the results of Frantz et al. (2004) show a comparable pattern. Even though obvious variation in the individual sampling frequency existed, almost all badgers present in the area were sampled. Thus, the rarefaction analysis yielded quite reliable results. However, the IH observed in the three badger studies might have been crucial in other populations or situations, e.g. when a lower proportion of the population is represented in the samples, especially when CR methods are applied (Pollock 1982). Furthermore, video control as well as direct observation in all three studies focused on obtaining an independent census of the sampled badger groups; not on observing the sampling behaviour itself. As a consequence, potential sources of IH such as dependence of latrine use or access to bait on social status may remain undetected. The same seems to hold true for a study on the lesser horseshoe bat *Rhinolophus hipposideros*; the non-invasive population estimate was validated via direct counts of bats in their day roosts (Puechmaile & Petit 2007). The direct counts did not reveal any IH in the faeces sampling. However, IH was detected in the sampling data of several of the sampled bat colonies via likelihood-ratio or simulation tests.

In a faeces sampling study on wolves *Canis lupus* in the Italian Alps, Marucco et al. (2009) used an evaluation system for age-dependent marking behaviour related to defecation. Due to the fact that part of the population was radio-collared or otherwise known, it was possible to discriminate between faeces deposited by adult wolves for marking purposes and 'non-marking' faeces. The authors detected age- and status dependent IH in the defecation behaviour and concluded that they would have missed a considerable part of the juvenile population, if they had not adapted their search pattern. However, by means of the field tests, Marucco et al. (2009) were able to apply and confirm a representative faeces sampling strategy.

Discussion

Most researchers seem to be aware of IH being a major problem present in population estimation based on non-invasive genetic methods. The vast majority of articles dealing with non-invasive methods applied for such purpose mention or discuss this problem. In most studies based on CR approaches, the authors attempted to account for

potential bias by employing models which incorporate IH (Chao 1987). However, as long as the different sources and the extent of IH are unknown, the results of population estimations are strongly model dependent and might not reflect reality (Link 2003, 2004). Furthermore, the different methods to test for IH in the data set may have limited power and thus often fail to detect IH (Boulanger & McLellan 2001, Miller et al. 2005). In the literature, it has been mentioned that the power of such test procedures is especially low for small p s (Menkens & Anderson 1988, Boulanger & McLellan 2001). The results of our analysis support this finding; they indicate an impact of p on the detectability of IH via GOF tests and model selection procedures. This effect does not seem to depend on the type of test and the software used (logistic regression with test type as additional covariate showed no significant effect; results not shown). In our analysis, we used a very conservative approach by averaging over the studies and using both weighted and unweighted values. Both approaches show very similar results, indicating that the results are robust to details in the analysis. In studies with a low p , IH was detected considerably less often than in studies with a higher p . The highest proportion of detected IH was attributed to studies with p between 0.2 and below 0.4. Interestingly, in none of the three studies reporting $p > 0.4$, tests suggested the incidence of IH. This might be due to the fact that IH bias becomes much less problematic, perhaps even negligible, when p is high, which has been shown in simulation studies (J. Boulanger, pers. comm.). When most animals in a population are actually captured or sampled, the differences in p between individuals have much less impact on the population estimate (Pollock 1990, Lukacs & Burnham 2005). Thus, IH might not be reflected in GOF testing or model selection when p is high. Even though the number of sampling occasions for each study was not included in the most supported model, there seems to be an indication of a certain influence on the detectability of IH, because the model including sampling occasions was ranked only marginally inferior to the best model. The more sampling occasions that are carried out, the better might be the ability to detect IH via GOF-testing. However, this point needs further investigation before a clear conclusion on sampling occasions can be drawn. In contrast to p , an impact of coverage on the detectability of IH was not supported in our analysis; despite the correlation of coverage and p .

It should be mentioned that with our analysis, we are not able to distinguish if a negative result of the testing for IH is due to lack of power and test failure or because there simply is no IH present in a given data set. However, regarding the existing literature including simulation studies and studies on populations of known size, IH seems to be almost ubiquitous in non-invasive sampling data sets like in the conventional CR (Pollock 1990, Borchers et al. 2002, Knapp et al. 2007, Lukacs & Burnham 2005). Therefore, it seems much more likely that a negative test result is caused by a low test power than that a data set is really homogeneous, particularly when p is low. In recent years, new modelling approaches, e.g. multistate and multi-event models (Pradel 2005), have been developed, which might allow a more flexible handling of CR data in presence of IH (Crespin et al. 2008). In this context GOF testing using non-parametric methods, like in program U-CARE, seems to be quite promising compared to conventional methods (Choquet et al. 2009, Cubaynes et al. 2010).

In five of the articles we reviewed, IH was not detected by the data tests which were performed. However, in two of the five cases, additional field tests were carried out, and both of them revealed IH. In general, data tests and/or pronounced differences in the 'capture frequencies' indicate presence of IH bias without carrying out extra field tests, but often further investigations would be required to uncover the causes of IH. Since many different IH sources exist, they can influence estimations in different ways, and this effect may also depend on the sampling design (Crespin et al. 2008). Models that are relatively robust to IH generally show reduced precision of estimate (Boulanger 2004a). This may not be tolerable in cases where an accurate estimate is particularly important, e.g. when the spread of diseases is concerned (Artois et al. 2002) or when management plans for rare or endangered species are considered (Guschanski et al. 2009). However, for endangered species, overestimating a population is much more critical than underestimating it (Meijer et al. 2008), so some underestimation bias may be tolerable in certain cases.

Assessment of IH via field tests

The choice of methods to test for IH in the field strongly depends on the observed species and its behavioural patterns, as well as its space and habitat use. Thus, e.g. for badgers which live in social groups, share setts and make rather small-scale

movements, video control or direct observations at setts seem to be an adequate method to validate non-invasively obtained estimates (see e.g. Frantz et al. 2004 and Scheppers et al. 2007). Contrastingly, for highly mobile species such as bears and coyotes, radio-telemetry may be more promising. The suitability of a field method to test for IH furthermore depends on the applied sampling strategy. For example, video control or direct observations can be appropriate for surveillance of discrete sampling stations like hair traps or badger setts, but will not be suitable for large-scale sampling designs such as e.g. line transects. Radio-telemetry may be more effective to observe movements and transect- or trap-encounter rates of animals on a large scale. Furthermore, radio-telemetry is useful for obtaining information on spatial distribution and home-range sizes of animals in order to fit sampling designs and to account for closure violations and edge effects (Boulanger et al. 2004a, Dreher et al. 2007). The feasibility of a sampling method for a given species or population can depend on spatial characteristics like home-range sizes and distribution of animals in the sampled area. Settlage et al. (2008) found hair sampling of black bears via baited sampling stations impractical for the Southern Appalachian region. Due to small home-range sizes of the resident bears, sampling probabilities were low and biased. In order to yield a reliable estimate, a much higher sampling intensity would have been necessary (Settlage et al. 2008). Grizzly bears showed considerably higher p with comparable sampling intensities because of their larger average home ranges (Boulanger et al. 2004a, McLoughlin et al. 2003).

Interactions between sampling strategies and study species' characteristics

The occurrence and/or extent of IH may differ dependent on the applied sampling method. 'Active' sampling methods like hair sampling via baited hair traps presuppose that animals actively approach the sampling station. In many species, it has been shown that individuals show consistent or context-specific personality traits, e.g. they differ in their exploration behaviour and their reactions towards newly introduced factors, which may affect their sampling probability (Coleman & Wilson 1998, Ruis et al. 2000, Dingemanse et al. 2003, Mettke-Hofmann et al. 2005). Furthermore individual experience and life history may influence behaviour with respect to sampling stations. This could cause IH which is not

necessarily related to sex, age or social status, and which might be hard to quantify and very difficult to account for in a model. Thus, in some cases, it may be reasonable to apply a different sampling method. In this context, 'passive' sampling strategies such as faeces sampling along transects represent an alternative, which may be less affected by individual behaviour or status differences. This may hold true particularly for group living species. Interactions between animals can increase IH, especially when sampling concentrates on defined stations such as hair traps, which require active approach. As an example, we conducted a hair sampling pilot study on wild boar *Sus scrofa*. Video observation at baited hair traps revealed significant behavioural differences depending on age of the animals and on their group status (Ebert et al. 2010). However, also for bears which can be considered as living mainly solitary, it has been shown that via faeces sampling, a 'passive' sampling strategy, a larger part of a population can be observed than at hair sampling as an 'active' approach (Wasser et al. 2004). Wasser et al. (2004) applied both methods in the same study area and time period, and via hair sampling only 46% of the individuals that were identified via faeces sampling were detected. 'Passive' sampling methods in most cases will not yield completely unbiased results (in fact, most of the faeces sampling studies that we reviewed reported IH in their data sets). Nevertheless, 'passive' sampling might rule out certain sources of IH which are not avoidable in 'active' approaches, and thus holds the potential to yield results with smaller overall bias. However, in some (especially social and/or territorial) species, status or age differences between individuals may cause differences in faeces deposition patterns leading to IH in detection probabilities. This has been shown e.g. for wolves (Marucco et al. 2009, Cubaynes et al. 2010). Thus, 'passive' sampling will not be suitable in all cases, and at any rate the appropriate sampling strategy and design have to be carefully tested for each particular species and population. Furthermore, the DNA quality of faeces in some cases has been shown to be inferior to that of hair, thus population estimates derived from faeces sampling data may be more in danger of bias due to genotyping errors (Piggott & Taylor 2003).

Recommendations

1) We recommend to perform a pilot study not only in the lab, but also in the field: In any case, it is most

advisable that researchers who plan to establish population estimation based on non-invasive genetic sampling perform pilot studies not only to assess genotyping error rates, but also to detect sources of IH bias in the field. The fact that the majority of reviewed studies in which such field tests were performed actually detected IH highly supports this recommendation. The appropriate methods to assess IH in the field depend on the species or population under study as well as on the applied sampling method;

2) not to rely solely on GOF testing and model selection procedures: This holds true, especially when p and coverage are low! It can be reasonable to incorporate heterogeneity in an estimation model, even if tests suggest that there is no IH present in a data set, because their power is often low. It is always recommended to include biological knowledge and information about e.g. study species and habitat to validate model choice;

3) try to reduce IH by adapting the sampling design: Knowledge about the sources and extent of IH can enable researchers to adapt the sampling design to account for the bias. Among the methods to reduce IH bias in the field, the application of two or more different sampling strategies in combination seems especially promising (Dreher et al. 2007, Boulanger et al. 2008, Settlage et al. 2008). If multiple methods are used simultaneously to sample a population, the impact of IH caused by any single method can be minimised (Pollock 1982, Williams et al. 2002). The improvement of estimations based on multiple approaches increases with decreasing correlation between the applied sampling methods (Boulanger et al. 2008). In the case of hair sampling, the use of unbaited sampling stations (e.g. installed at trails or rubbing trees) and changing of sampling locations between sessions may be applied to reduce IH due to competition between individuals for resources and due to 'trap happy' individuals (Scheppers et al. 2007, Boulanger et al. 2008). Collection of faeces samples with the help of trained dogs seemed to increase the detection rate and efficiency of the method considerably, allowing a relatively representative and unbiased population survey compared e.g. to hair sampling (Wasser et al. 2004, Long et al. 2007). Furthermore, it is advisable to perform a sufficiently high number of sampling occasions in order to increase the overall sampling probability and thus to facilitate accurate estimates;

4) try to sample a large part of the population: As shown e.g. in the three badger studies we reviewed,

one effective way to reduce the bias caused by IH is to sample a large proportion of the population (Lukacs & Burnham 2005). This is generally desirable and has been recommended in relevant literature many times before (see e.g. Otis et al. 1978, Pollock et al. 1982), but is certainly not always feasible. Furthermore, an increase in sample size can have an unfavourable impact on non-invasive genetic population estimates; the more samples are analysed, the higher the misidentification rate due to genotyping errors (McKelvey & Schwartz 2004). Thus, careful error-checking protocols for genotyping are crucial and genotyping error rates should be determined in order to avoid an increase in bias through misidentification (Maudet et al. 2004, Roon et al. 2005b);

5) consider switching to other sampling strategies: Adapting the sampling design may not always be possible or may yield no success. Furthermore, an unsolved problem still remains; even though detection of IH and its sources may be possible with methods like e.g. radio-telemetry or video observation, the exact quantification of such variation and thus its incorporation in estimation models still seems to be very difficult. Consequently, when reduction and/or modelling of IH is not possible, it can be recommendable to apply a different sampling method in some cases. The suitability of a method can depend e.g. on characteristics of the studied species, population or study area. In some cases, 'passive' sampling approaches may yield more representative results than 'active' methods. In case IH cannot be reduced or avoided, a study should be designed in such a way that it results in capture probabilities between 0.2 and 0.4, to have an ample chance to detect existing IH.

In conjunction with problems caused by genotyping errors, IH is a highly challenging issue in non-invasive population estimation. It is a well-known and explicitly discussed problem at least with regard to its theoretical and model-based aspects. IH can be identified and strongly reduced when field sampling design and analytical approach are carefully prepared. However, more attention should be given to the evaluation of field methods to bring forward more effective and sustainable population estimates, which is especially important for conservation of endangered species and even more in fragmented habitats.

Acknowledgements - we wish to thank J. Boulanger for giving valuable supplemental ideas and J. Arnold, J.

Hofmann, W. Maurer, K. Jochum, as well as two anonymous reviewers, for providing further helpful comments on earlier drafts of this manuscript. Our study was granted by the foundation 'Rheinland-Pfalz fuer Innovation' and the Ministry for Environment, Forestry and Consumer Protection in Rhineland-Palatinate, Germany. C. Ebert gratefully acknowledges financial support from the FAZIT foundation.

References

- Artois, M., Depner, K.R., Guberti, V., Hars, J., Rossi, S. & Rutili, D. 2002: Classical swine fever (*hog cholera*) in wild boar in Europe. - *Revue Scientifique et Technique de l'Office International des Epizooties* 21: 287-303.
- Baber, D.W. & Coblenz, B.E. 1986: Density, home range, habitat use, and reproduction in feral pigs on Santa Catalina Island. - *Journal of Mammalogy* 67: 512-525.
- Banks, S.C., Piggott, M.P., Hansen, B.D., Robinson, N.A. & Taylor, A.C. 2002: Wombat coprogenetics: enumerating a common wombat population by microsatellite analysis of faecal DNA. - *Australian Journal of Zoology* 59: 193-214.
- Banks, S.C., Hoyle, S.D., Horsup, A., Sunnucks, P. & Taylor, A.C. 2003: Demographic monitoring of an entire species (the northern hairy-nosed wombat, *Lasiorhinus kreffii*) by genetic analysis of non-invasively collected material. - *Animal Conservation* 6: 101-107.
- Bates, D. & Maechler, M. 2010: lme4: Linear mixed-effects models using Eigen and S4 classes. R package version 0.999375-33. - Available at: <http://CRAN.R-project.org/package=lme4> (Last accessed on 15 June 2010).
- Belant, J.L., Van Stappen, J.F. & Paetkau, D. 2005: American black bear population size and genetic diversity at Apostle Islands National Lakeshore. - *Ursus* 16: 85-92.
- Bellemain, E., Nawaz, M.A., Valentini, A., Swenson, J.E. & Taberlet, P. 2007: Genetic tracking of brown bear in northern Pakistan and implications for conservation. - *Biological Conservation* 134: 537-547.
- Bellemain, E., Swenson, J.E., Tallmon, D., Brunberg, S. & Taberlet, P. 2005: Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for brown bears. - *Conservation Biology* 19: 150-161.
- Boersen, M.R., Clark, J.D. & King, T.L. 2003: Estimating black bear population density and genetic diversity at Tensas River, Louisiana using microsatellite DNA markers. - *Wildlife Society Bulletin* 31: 197-207.
- Borchers, D.L., Buckland, S.T. & Zucchini, W. 2002: Estimating Animal Abundance - Closed Populations. - *Statistics for Biology and Health*, Springer, London, UK, 314 pp.
- Boulanger, J., Himmer, S. & Swan, C. 2004a: Monitoring of grizzly bear population trends and demography using DNA capture-recapture methods in the Owikeno Lake

- area of British Columbia. - *Canadian Journal of Zoology* 82: 1267-1277.
- Boulanger, J., Kendall, K.C., Stetz, J.C., Roon, D.A., Waits, L.P. & Paetkau, D. 2008: Multiple data sources improve DNA-based capture-recapture estimates of grizzly bears. - *Ecological Applications* 18: 577-589.
- Boulanger, J. & McLellan, B. 2001: Closure violation in DNA-based capture-recapture estimation of grizzly bear populations. - *Canadian Journal of Zoology* 79: 642-651.
- Boulanger, J., McLellan, B.N., Woods, J.G., Proctor, M.F. & Strobeck, C. 2004b: Sampling design and bias in DNA-based capture-recapture population and density estimates of grizzly bears. - *Journal of Wildlife Management* 68: 457-469.
- Boulanger, J., Proctor, M., Himmer, S., Stenhouse, G., Paetkau, D. & Cranston, J. 2006: An empirical test for DNA capture-recapture sampling strategies for grizzly bears. - *Ursus* 17: 149-158.
- Boulanger, J., Stenhouse, G. & Munro, R. 2004c: Sources of heterogeneity bias when DNA capture-recapture sampling methods are applied to grizzly bear (*Ursus arctos*) populations. - *Journal of Mammalogy* 85: 618-624.
- Boulanger, J., White, G.C., McLellan, B., Woods, J., Proctor, M. & Himmer, S. 2002: A meta-analysis of grizzly bear DNA capture-recapture projects in British Columbia, Canada. - *Ursus* 13: 137-152.
- Broquet, T., Ménard, N. & Petit, E. 2007: Non-invasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. - *Conservation Genetics* 8: 249-260.
- Broquet, T. & Petit, E. 2004: Quantifying genotyping errors in non-invasive population genetics. - *Molecular Ecology* 13: 3601-3608.
- Burnham, K.P. & Anderson, D.R. 1998: Model selection and multimodel inference - a practical information-theoretic approach. - Springer Verlag, 488 pp.
- Chao, A. 1987: Estimating the population size for capture-recapture data with unequal catchability. - *Biometrics* 43: 783-791.
- Chao, A.L. & Jeng, S.L. 1992: Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal. - *Biometrics* 48: 201-216.
- Choquet, R., Lebreton, J.-D., Gimenez, O., Reboulet, A.-M. & Pradel, R. 2009: U-CARE: Utilities for performing goodness of fit tests and manipulating CAPTURE-RECAPTURE data. - *Ecography* 32: 1071-1074.
- Coleman, K. & Wilson, D.S. 1998: Shyness and boldness in pumpkinseed sunfish: individual differences are context-specific. - *Animal Behaviour* 56: 927-936.
- Creel, S., Spong, G., Sands, J.L., Rotella, J., Zeigler, J., Joe, L., Murphy, K.M. & Smith, D. 2003: Population size estimation in Yellowstone wolves with error-prone non-invasive microsatellite genotypes. - *Molecular Ecology* 12: 2003-2009.
- Crespin, L., Choquet, R., Lima, M., Merritt, J. & Pradel, R. 2008: Is heterogeneity of catchability in capture-recapture studies a mere sampling artefact or a biologically relevant feature of the population? - *Population Ecology* 50: 247-256.
- Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J.-M., Lebreton, J.-D., Marboutin, E., Miquel, C., Reboulet, A.-M., Poillot, C., Taberlet, P. & Gimenez, O. 2010: Importance of accounting for detection heterogeneity when estimating abundance: the case of French wolves. - *Conservation Biology* 24: 621-626.
- Dingemans, N.J., Both, C., van Noordwijk, A.J., Rutten, A.L. & Drent, P. 2003: Natal dispersal and personalities in great tits (*Parus major*). - *Proceedings of the Royal Society (London)* 270: 741-747.
- Dreher, B.P., Winterstein, S.R., Scribner, K.T., Lukacs, P.M., Etter, D.R., Rosa, G.J.M., Lopez, V.A., Libants, S. & Filcek, K.B. 2007: Non-invasive estimation of black bear abundance incorporating genotyping errors and harvested bear. - *Journal of Wildlife Management* 71: 2684-2693.
- Ebert, C., Huckschlag, D., Schulz, H.K. & Hohmann, U. 2010: Can hair traps sample wild boar (*Sus scrofa*) randomly for the purpose of non-invasive population estimation? - *European Journal of Wildlife Research* 56: 583-590.
- Eggert, L.S., Eggert, J.A. & Woodruff, D.S. 2003: Estimating population size for elusive animals: the forest elephants of Kakum National Park, Ghana. - *Molecular Ecology* 12: 1389-1402.
- Frantz, A.C., Pope, L.C., Carpenter, P.J., Roper, T.J., Wilson, G.J. & Delahay, R.J. 2003: Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. - *Molecular Ecology* 12: 1649-1661.
- Frantz, A.C., Schaul, M., Pope, L.C., Fack, F., Schley, L., Muller, C.P. & Roper, T.J. 2004: Estimating population size by genotyping remotely plucked hair: the Eurasian badger. - *Journal of Applied Ecology* 41: 985-995.
- Gervasi, V., Ciucci, P., Boulanger, J., Posillico, M., Sulli, C., Focardi, S., Randi, E. & Boitani, L. 2008: A preliminary estimate of the Apennine brown bear population size based on hair-snag sampling and multiple data source capture-recapture Huggins models. - *Ursus* 19: 105-121.
- Guschanski, K., Vigilant, L., McNeilage, A., Gray, M., Kagoda, E. & Robbins, M.M. 2009: Counting elusive animals: Comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. - *Animal Conservation* 14: 290-300.
- Ihaka, R. & Gentleman, R. 1996: R: A language for data analysis and graphics. - *Journal of Computational and Graphical Statistics* 5: 299-314.
- Immell, D. & Anthony, R.G. 2006: Estimation of black bear abundance using a discrete DNA sampling device. - *Journal of Wildlife Management* 72: 324-330.

- Jacob, G., Debrunner, R., Gugerli, F., Schmid, B. & Bollmann, K. 2010: Field surveys of capercaillie (*Tetrao urogallus*) in the Swiss Alps underestimated local abundance of the species as revealed by genetic analyses of non-invasive samples. - *Conservation Genetics* 11: 33-44.
- Knapp, S.H., Craig, B.A. & Waits, L.P. 2007: Incorporating genotyping error into non-invasive DNA-based capture-recapture population estimates. - *Journal of Wildlife Management* 73: 598-604.
- Kohn, M.H., York, E.C., Kamradt, D.A., Haught, G., Sauvajot, R.M. & Wayne, R.K. 1999: Estimating population size by genotyping faeces. - *Proceedings of the Royal Society (London)* 266: 657-663.
- Link, W.A. 2003: Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. - *Biometrics* 59: 1123-1130.
- Link, W.A. 2004: Individual heterogeneity and identifiability in capture-recapture models. - *Animal Biodiversity and Conservation* 27: 87-91.
- Long, R.A., Donovan, T.M., Mackay, P., Zielinski, W.J. & Buzas, J.S. 2007: Comparing faeces detection dogs, cameras, and hair snares for surveying carnivores. - *Journal of Wildlife Management* 71: 2018-2025.
- Lukacs, P.M. & Burnham, K.P. 2005: Review of capture-recapture methods applicable to non-invasive genetic sampling. - *Molecular Ecology* 14: 3909-3919.
- Marucco, F., Pletscher, D.H., Boitani, L., Schwartz, M.K., Pilgrim, K.L. & Lebreton, J.-D. 2009: Wolf survival and population trend using non-invasive capture-recapture techniques in the Western Alps. - *Journal of Applied Ecology* 46: 1003-1010.
- Maudet, C., Luikart, G., Dubray, D., von Hardenberg, A. & Taberlet, P. 2004: Low genotyping error rates in wild ungulate faeces sampled in winter. - *Molecular Ecology Notes* 4: 772-775.
- McKelvey, K.S. & Pearson, D.E. 2001: Population estimation with sparse data: the role of estimators versus indices revisited. - *Canadian Journal of Zoology* 79: 1754-1765.
- McKelvey, K.S. & Schwartz, M.K. 2004: Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. - *Journal of Wildlife Management* 68: 439-448.
- McLoughlin, P.D., Cluff, H.D., Gau, R.J., Mulders, R., Case, R. & Messier, F. 2003: Effect of spatial differences in habitat on home ranges of grizzly bears. - *Écoscience* 10: 11-16.
- Meijer, T., Norén, K., Hellström, P., Dalén, L. & Angerbjörn, A. 2008: Estimating population parameters in a threatened arctic fox population using molecular tracking and traditional field methods. - *Animal Conservation* 11: 330-338.
- Menkens, G.E. & Anderson, S.H. 1988: Estimation of small-mammal population size. - *Ecology* 69: 1952-1959.
- Mettke-Hofmann, C., Ebert, C., Schmidt, T., Steiger, S. & Stieb, S. 2005: Personality traits in resident and migratory warbler species. - *Behaviour* 142: 1357-1375.
- Miller, C.R., Joyce, P. & Waits, L.P. 2005: A new method for estimating the size of small populations from genetic capture-recapture data. - *Molecular Ecology* 14: 1991-2005.
- Mills, L.S., Citta, J.J., Lair, K.P., Schwartz, M.K. & Tallmon, D.A. 2000: Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. - *Ecological Applications* 10: 283-294.
- Minta, S. & Mangel, M. 1989: A simple population estimate based on simulation for capture-recapture and capture-resight data. - *Ecology* 70: 1738-1751.
- Miquel, C., Bellemain, E., Poillot, C., Bessière, J., Durand, A. & Taberlet, P. 2006: Quality indexes to assess the reliability of genotypes in studies using non-invasive sampling and multiple-tube approach. - *Molecular Ecology Notes* 6: 985-988.
- Mowat, G., Heard, D.C., Seip, D.R., Poole, K.G., Stenhouse, G. & Paetkau, D.W. 2005: Grizzly *Ursus arctos* and black bear *Ursus americanus* densities in the interior mountains of North America. - *Wildlife Biology* 11(1): 31-48.
- Mowat, G. & Paetkau, D. 2002: Estimating marten *Martes americana* population size using hair capture and genetic tagging. - *Wildlife Biology* 8(3): 201-209.
- Mowat, G. & Strobeck, C. 2000: Estimating population size of Grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. - *Journal of Wildlife Management* 64: 183-193.
- Obbard, M.E., Howe, E.J. & Kyle, C.J. 2010: Empirical comparison of density estimators for large carnivores. - *Journal of Applied Ecology* 47: 76-84.
- Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R. 1978: Statistical inference for capture-recapture experiments. - *Wildlife Monographs* 62: 1-135.
- Paetkau, D. 2003: An empirical exploration of data quality in DNA-based population inventories. - *Molecular Ecology* 12: 1375-1387.
- Petit, E. & Valière, N. 2006: Estimating population size with non-invasive capture-recapture data. - *Conservation Biology* 20: 1062-1073.
- Piggott, M.P. & Taylor, A.C. 2003: Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. - *Wildlife Research* 30: 1-13.
- Pledger, S. & Efford, M. 1998: Correction of bias due to heterogeneous capture probability in capture-recapture studies in open populations. - *Biometrics* 54: 888-898.
- Pollock, K.H., Nichols, J.D., Brownie, C. & Hines, J.E. 1990: Statistical inference for capture-recapture experiments. - *Wildlife Monographs* 107: 1-97.
- Pollock, K.H., Nichols, J.D., Brownie, C. & Hines, J.E. 1982: Statistical inference for capture-recapture experiments. - *Wildlife Monographs* 107: 1-97.
- Poole, K.G., Mowat, G. & Fear, D.A. 2001: DNA-based

- population estimate for grizzly bears *Ursus arctos* in northeastern British Columbia, Canada. - *Wildlife Biology* 7(2): 105-115.
- Pradel, R. 2005: Multievent, an extension of multistate capture-recapture models to uncertain states. - *Biometrics* 61: 442-447.
- Prigioni, C., Remonti, L., Balestrieri, A., Sgroso, S., Priore, G., Mucci, N. & Randi, E. 2006: Estimation of European otter (*Lutra lutra*) population size by fecal DNA typing in southern Italy. - *Journal of Mammalogy* 87: 855-858.
- Prugh, L.R., Ritland, C.E., Arthur, S.M. & Krebs, C.J. 2005: Monitoring coyote population dynamics by genotyping faeces. - *Molecular Ecology* 14: 1585-1596.
- Puechmaile, S.J. & Petit, E. 2007: Empirical evaluation of non-invasive capture-capture-recapture estimation of population size based on a single sampling session. - *Journal of Applied Ecology* 44: 843-852.
- Roon, D.A., Thomas, M.E., Kendall, K.C. & Waits, L.P. 2005a: Evaluating mixed samples as a source of error in non-invasive genetic studies using microsatellites. - *Molecular Ecology* 14: 195-201.
- Roon, D.A., Waits, L.P. & Kendall, K.C. 2005b: A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. - *Animal Conservation* 8: 203-215.
- Rudnick, J.A., Katzner, T.E., Bragin, E.A. & DeWoody, J.A. 2008: A non-invasive genetic evaluation of population size, natal philopatry, and roosting behaviour of non-breeding eastern imperial eagles (*Aquila heliaca*) in central Asia. - *Conservation Genetics* 9: 667-676.
- Ruis, M.A.W., te Brake, J.H.A., van de Burgwal, J.A., de Jong, I.C., Blokhuis, H.J. & Koolhaas, J. 2000: Personalities in female domesticated pigs: behavioural and physiological indications. - *Applied Animal Behavioural Sciences* 66: 31-47.
- Schepers, T.L.J., Frantz, A.C., Schaul, M., Engel, E., Breyne, P., Schley, L. & Roper, T.J. 2007: Estimating social group size of Eurasian badgers *Meles meles* by genotyping remotely plucked single hairs. - *Wildlife Biology* 13(2): 195-207.
- Settlage, K.E., Van Manen, F.T., Clark, J.D. & King, T.L. 2008: Challenges of DNA-based capture-recapture studies of American black bears. - *Journal of Wildlife Management* 72: 1035-1042.
- Solberg, K.H., Bellemain, E., Drageset, O-M., Taberlet, P. & Swenson, J.E. 2006: An evaluation of field and non-invasive genetic methods to estimate brown bear (*Ursus arctos*) population size. - *Biological Conservation* 128: 158-168.
- Sweitzer, R., van Vuren, D., Gardner, I.A., Boyce, W. & Waithman, J.D. 2000: Estimating sizes of wild pig populations in the north and central coast regions of California. - *Journal of Wildlife Management* 64: 531-543.
- Taberlet, P., Waits, L.P. & Luikart, G. 1999: Non-invasive genetic sampling: look before you leap. - *Trends in Ecology and Evolution* 14: 323-327.
- Triant, D.A., Pace, R.M., III. & Stine, M. 2004: Abundance, genetic diversity and conservation of Louisiana black bears (*Ursus americanus luteolus*) as detected through non-invasive sampling. - *Conservation Genetics* 5: 647-659.
- Wasser, S.K., Davenport, B., Ramage, E.R., Hunt, K.E., Parker, M., Clarke, C. & Stenhouse, G. 2004: Faeces detection dogs in wildlife research and management: application to grizzly and black bears in the Yellowhead Ecosystem, Alberta, Canada. - *Canadian Journal of Zoology* 82: 475-492.
- White, G.C., Andersen, D.R., Burnham, K.P. & Otis, D.L. 1982: Capture-recapture and removal methods for sampling closed populations. - Los Alamos National Laboratory Report LA-8787-NERP, Los Alamos, New Mexico, USA, 235 pp.
- White, G.C. & Burnham, K.P. 1999: Program MARK: survival estimation from populations of marked animals. - *Bird Study* 46: 120-138.
- Williams, B.K., Nichols, J.D. & Conroy, M.J. 2002: Analysis and management of animal populations. - Academic Press, San Diego, California, USA, 817 pp.
- Wilson, G.J., Frantz, A.C., Pope, L.C., Roper, T.J., Burke, T.A., Cheeseman, C.L. & Delahay, R.J. 2003: Estimation of badger abundance using faecal DNA typing. - *Journal of Applied Ecology* 40: 658-666.
- Woods, J.G., Paetkau, D., Lewis, D., McLellan, B.N., Proctor, M. & Strobeck, C. 1999: Genetic tagging of free-ranging black and brown bears. - *Wildlife Society Bulletin* 27: 616-627.