



Which sharks attract research? Analyses of the distribution of research effort in sharks reveal significant non-random knowledge biases

Simon Ducatez 

Received: 23 March 2018 / Accepted: 8 March 2019 / Published online: 29 March 2019
© Springer Nature Switzerland AG 2019

Abstract Research effort is unevenly distributed across species, which can cause important biases in our understanding of evolutionary and ecological processes and affect conservation decisions. For example, many shark species remain understudied, despite the fact that sharks play fundamental roles in marine ecosystems and are particularly affected by fisheries. Assessing and acknowledging these differences in research effort across shark species is a key step to improving our knowledge and management of shark populations as it allows researchers to both target species in need of research and control for potential biases when performing comparative analyses. I provide here an index of research effort (the number of articles published between 1978 and 2014 listed in the Zoological Record database) for 509 shark

species that can be downloaded and utilized in future comparative analyses. I then show that research effort is not randomly distributed across species within the clade, but is significantly predicted by taxonomy, geography, ecology, life history and extinction risk. For example, large species that occur across a large range of latitudes and at shallow minimum depths have attracted considerably more research. The consequence is a strong bias in our knowledge of sharks towards species that are not an unbiased representative sample of the clade. The database and the patterns demonstrated here call for more awareness of the biases in research effort and their potential consequences, not only for practical considerations such as the interpretation of the results of comparative analyses and the assessment of extinction risk, but also for our basic understanding of the ecological roles of sharks.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11160-019-09556-0>) contains supplementary material, which is available to authorized users.

S. Ducatez (✉)
School of Biological Sciences, The University of Sydney,
Sydney, Australia
e-mail: simon.ducatez@gmail.com

S. Ducatez
Department of Biology, McGill University, Montréal,
Canada

S. Ducatez
CREAF, Cerdanyola del Vallès, Catalonia, Spain

Keywords Chondrichthyes · Extinction risk · Research bias · Research effort · Shark

Introduction

Sharks are a particularly iconic group of Chondrichthyes, studied predominantly because of their importance in fisheries (Barker and Schluessel 2005; Molina and Cooke 2012; Worm et al. 2013; Dulvy

et al. 2014) and their central role in marine ecosystems (Heithaus et al. 2008; Ferretti et al. 2010; Roff et al. 2016). The slow life history of most shark species, with late age-at-maturity and low fecundity, limits their ability to sustain fishing effort (Smith et al. 1998; Simpfendorfer and Kyne 2009). As a consequence, assessing the sustainability of shark populations, as well as of shark fisheries, is a key concern (Dulvy et al. 2014). Recent massive efforts from the scientific community allowed extinction risk assessments to be completed for over 260 shark species (Dulvy et al. 2014; IUCN 2017). Still, extinction risk remains unknown for about 250 described species, revealing the existence of important knowledge gaps within this group and suggesting that research efforts have not been evenly distributed across shark species. Awareness of such knowledge gaps is necessary in order to determine whether and how to draw general conclusions from existing data, and to improve both general knowledge and the efficiency of conservation actions. The key roles of sharks within both ecosystems and human activities makes identifying such gaps exceptionally important (Barker and Schluessel 2005; Ferretti et al. 2010; Dulvy et al. 2014).

Knowledge gaps created by interspecific variation in research effort within a taxon can complicate the interpretation of studies aimed at establishing general patterns and mechanisms, such as reviews, meta-analyses and comparative analyses. Such gaps can have important consequences for four main reasons. First, differences in research effort can directly bias species trait estimates. For example, estimates of maximum longevity will likely depend on the amount of research conducted on a given species: extreme values are more likely to be detected in heavily studied taxa (e.g. see Ducatez and Lefebvre 2014). Second, differences in knowledge across species will affect the accuracy of the traits measured (Garamszegi and Møller 2010, 2011, 2012). For example, if 100 individuals are used to assess the average value of a trait (e.g. body size at maturity, litter size...) for a given species, but only five individuals for another species, the accuracy of the measurement will likely be better for the first species. Third, because of knowledge gaps (e.g. because the value of a trait of interest is not known), some species have to be removed from analyses, even those that are intended to be global analyses of all shark species. For example, a study of the evolution of shark reproductive mode will

be focused on species in which the reproductive mode is known (about 300 species), ignoring a significant number of species (about 200) for which this trait is unknown. This raises concerns with regards to how representative the analyses can be, as removing less known species may bias the conclusions. Finally, these knowledge gaps may affect conservation issues, and especially extinction risk assessments (e.g. for the IUCN red list). These assessments rely on species knowledge to estimate how likely a given species is to become extinct. Differences in species knowledge may lead to differences in the accuracy of the risk estimations, while species that are poorly known are generally not assessed at all (or considered as “data deficient”). Understudied groups must therefore be identified to ensure that risk assessment procedures are both possible and unbiased.

To account for or circumvent these limitations, we first need to operationally measure these knowledge differences. The number of scientific publications produced per species has been frequently used in the literature as a proxy for research effort (Sol et al. 2005; Brito and Oprea 2009; Arnold and Nunn 2010; de Lima et al. 2011; González-Suárez et al. 2012; Ducatez and Lefebvre 2014; Ducatez and Shine 2017). This metric allows for the identification of biases in research effort that are associated with species traits or taxonomy, allowing us to determine whether and how these biases may affect the conclusions of interspecific comparative studies. Ultimately, these indices can be used to correct for these biases, e.g. by weighing species based on how well investigated they are, or by directly correcting trait values by research effort (Garamszegi and Møller 2010, 2011, 2012). In addition, the scientific community can attempt to correct these discrepancies in species knowledge by directing new research towards species or groups that are poorly known.

Different scenarios may explain why research effort is not evenly distributed across species. Some species may be targeted by researchers because they have original characteristics that raised researchers' interest, such as a unique reproductive mode, or because their behaviour has consequences for human populations, e.g. shark species causing human injuries. For some traits, the causality may be indirect, for example if a characteristic targeted by researchers (e.g. involvement in fatal attacks on humans) is correlated with other traits (e.g. body size). Other

species may be easier to study, e.g. because they have a larger distribution, are larger or occur at shallower depths. Although biases in research effort have been investigated and included in a number of comparative analyses in birds and mammals, they have often been neglected for other groups, including sharks (but see Rasmussen and Randhawa 2018). The objectives of this study are thus to 1) provide an index of research effort for all described extant shark species (super-order Selachimorpha) and 2) assess the distribution of this measure across species, according to species traits, taxonomy and extinction risk.

The index of research effort for all extant shark species is freely available in the Supplementary Material (Table S1), to be used either to control for research effort as a confounding variable in future comparative analyses, or as a guideline to orientate future research on sharks. I also used this database to test whether research effort in sharks was associated with taxonomy, geographic region and twelve life history and ecological traits: minimum depth, maximum depth, depth range, habitat breadth, latitudinal range, trophic level, size at birth, maximum size, male and female age at maturity, maximum longevity and reproductive mode. Finally, I tested whether differences in research effort were associated with an extinction vulnerability metric.

Methods

Research effort estimation

Research effort cannot be directly and concretely measured but can be estimated using different indices. Importantly, these indices are likely subject to biases, and testing whether different potential metrics provide similar relative estimates of research effort is important to establish the validity of the indices considered. Here, to estimate global research effort for each of the 509 shark species included in Weigmann (2016), I built an index based on papers extracted from the Zoological Record database. This database is one of the most exhaustive compilations of the zoological literature and has been used to estimate research effort in other taxa (e.g. in birds, Ducatez and Lefebvre 2014). For each species, I extracted the number of publications referenced in this database between 1978 and 2014. The entire data set is presented as Table S1.

I used the Latin names included in Weigmann's checklist as reference names for each species, and made searches on keywords, abstracts and titles (using the search option by "topic"). To test whether this search method could in itself bias the index, I re-estimated research effort using article titles only, and compared this estimate with the previous one. Considering only titles or more components of articles to look for species names should cause different biases. A species name sometimes does not appear in the title of an article, even if the article is focused on that species. Conversely, species names sometimes are used in abstracts or keywords of articles even if the species is not investigated in the article (e.g. because it is used for comparison, as an example, or as the model species for the subject of the article). By comparing the two indices, I thus test the sensitivity of the index used as a proxy of research effort to the search method. As species and genus names are regularly modified with advances in taxonomy and some species are known in the literature by different names, their research effort is likely to be underestimated when considering only one name. I therefore collected the different Latin names known for each shark species by considering the names used in Compagno et al. (2005) and the list of synonyms provided on the IUCN web site (IUCN 2017). Though these references do not provide exhaustive lists of species synonyms, they allow to increase the accuracy of research effort estimates, and to test whether and how missing synonyms may affect the estimation of research effort. I reassessed research effort considering the different Latin names collected for each species. For example, the number of articles published on the great white shark was estimated by the number of references provided by Zoological Record when searching for: ("Carcharodon carcharias" OR "Squalus carcharias") in publications "topic" from 1978 to 2014. I then estimated the correlation between this new estimate of research effort and the one based on the nomenclature from Weigmann (2016)'s list only. The latter analysis will bring information on how missing some synonyms may affect research effort estimates. Note that shark taxonomy is regularly updated, and several changes have been published subsequent to the taxonomy of Weigmann (2016) used in this study. Many of these changes are listed by Weigmann (2017).

Geographic region

Regional differences in research effort frequently occur (e.g. Ducatez and Lefebvre 2014), often due to a nonrandom distribution of researchers and the difficulties involved in reaching or working in certain areas. I obtained data on species geographic region from Weigmann (2016), who used the classification of the International Hydrographic Organization (IHO 1953) to identify ten different areas. I included species occurring in more than one area in a separate category.

Traits

Ecological factors are likely to affect how easy a species is to study, and thus affect research effort. The depth (range, minimum and maximum depths), range of latitudes, and the diversity of habitats in which a species occurs may all influence how frequently it is studied. I extracted minimum and maximum depth, as well as depth range, from Weigmann (2016), and latitudinal range from Randhawa et al. (2015). To estimate habitat breadth, I considered the diversity of habitats in which a species is known to occur. I used the IUCN data on species habitat use and followed the habitat classification from Dulvy et al. (2014). I thus considered five different habitat types (coastal and continental shelf; pelagic; meso- and bathy-pelagic; deepwater; freshwater) and used the number of categories (from one to five) in which each species occurs as proxy of habitat breadth.

Traits such as trophic level, body size and life history traits might also relate to research effort, both because of biases in the choice of species for research projects, and because the trait estimates themselves may change as knowledge on a species increases (e.g. maximum longevity or maximum body size). I extracted trophic levels from Randhawa et al. (2015) and FishBase (Froese and Pauly 2017), maximum body size from Weigmann (2016), and life history data from the IUCN web site (IUCN 2017), including data on size at birth, male and female age at maturity, maximum longevity, and reproductive mode. For reproductive mode I used a binary variable that segregated species giving birth to free-swimming individuals from species that lay eggs.

Analyses

I first tested whether research effort was associated with each factor separately. This univariate approach allowed for the identification of all associations between research effort and trait values that should be acknowledged, even if these associations are driven by other correlated traits (e.g. if changes in research effort associated with size at birth are actually due to an effect of maximum body size). The subset of species for which the values of a given trait is known varied significantly between traits, and this univariate approach also allowed me to maximize the number of species included in each analysis. Following these individual analyses, I built a global model that included all significant predictors together in the same analysis.

Taxonomic biases

To determine which taxonomic level explains the highest proportion of the variance in research effort, I classified each species by its genus, family, and order. I then used linear mixed effects models to estimate the proportion of variance explained by each taxonomic level, using the lme procedure from the nlme R package. I included research effort (log-transformed) as response variable, and order, family and genus as nested random effects. I compared the AIC of the model without any higher taxonomic levels, versus the model including only order, versus the model with family and order, versus the model with genus, family and order. I estimated the proportion of variance explained by each taxonomic level calculating intra-class coefficients (ICC) for each of the three taxonomic levels, using variance estimates from the complete model (i.e. with the three taxonomic groups). I used maximum likelihood to compare the AIC for different models, but restricted maximum likelihood to get variance estimates used to calculate ICC.

Geographic biases

I tested whether the geographic region where species occur predicts research effort by including the geographic region in a linear model with research effort (log-transformed) as the response variable.

Traits

I tested whether research effort was associated with six ecological traits (minimum depth, maximum depth, depth range, habitat breadth, latitudinal range, trophic level) and six life history traits (maximum body size, size at birth, age at first reproduction in males and females, maximum longevity and reproductive mode). For continuous variables, I used Spearman correlation tests to check for correlations between these traits and research effort, as some of the variables did not follow a normal distribution. I used a Wilcoxon test for reproductive mode.

Some ecological and life history traits covary across species, such that, ultimately, one or a few traits might be the main drivers of all correlations found between research effort and the tested variables. To test for this possibility, I then adopted a multivariate approach, including all the predictors of research effort in a single linear mixed model (LMM) with research effort (log-transformed) as response variable and, as explanatory variables, the traits found to be significantly correlated with research effort (including taxonomy as a random effect, see details in the results section). These analyses allowed me to test whether taxonomic and species trait biases were confounded, as trait values might not be randomly distributed according to taxonomy, and as the different traits are likely to be inter-dependent. It also allowed me to determine which of the predictors of research effort in sharks were the most important.

Extinction risk

As research effort biases may have strong implications for conservation decisions, I also tested whether species extinction risk was associated with research effort. I used the IUCN Red List status as a measure of a species extinction risk, and after excluding Data Deficient and non-assessed species, I included the extinction risk in a linear model with research effort (log-transformed) as the response variable.

Results

Among shark species, the spiny dogfish *Squalus acanthias* had the highest research effort with a total of 739 articles (288 when restricting the search on

articles titles). It was followed by the blue shark (*Prionace glauca*, 448 articles), the great white shark (*Carcharodon carcharias*, 412 articles) and the small spotted catshark (*Scyliorhinus canicula*, 407 articles). Large iconic species such as the shortfin mako (*Isurus oxyrinchus*), scalloped hammerhead (*Sphyrna lewini*), basking shark (*Cetorhinus maximus*), whale shark (*Rhincodon typus*), bull shark (*Carcharhinus leucas*) or tiger shark (*Galeocerdo cuvier*) also were within the 15 most investigated shark species. In contrast, 41 species had a research effort of zero articles, including several catshark and lantern shark species (see Table S1).

Estimating research effort from searches made within titles, keywords and abstracts, or exclusively within titles, gave similar relative estimates: the two research effort estimates were highly correlated (Pearson correlation coefficient between the log-transformed research effort estimates = 0.923; $p = 0.001$). Changes in taxonomic nomenclature also had little effect on relative differences in research effort estimates. The research effort that considers multiple species names was highly correlated with the research effort considering one name only (Pearson correlation on log-transformed research efforts = 0.977; $p = 0.001$). All analyses detailed below considered research effort based on the most inclusive search parameters: multiple species names searched within titles, keywords and abstracts.

Taxonomy

The models including either family nested in order (AIC = 1655.6) or all three taxonomic levels (AIC = 1657.5) had a lower AIC than the null model (AIC = 1766.8) or the model including the order only (AIC = 1742.1). Considering either of these two best models, taxonomy explained about 40% of the variance in research effort (with order explaining 21 to 27% and family 16 to 23%, see Table 1). Research efforts within each order and family are given in Table 2 (see also Fig. 1).

Geographic biases

Geographic region also significantly explained variation in research effort ($F_{9,499} = 21.384$, $p < 0.001$). Species from the North Eastern Atlantic and Pacific Oceans and species occurring in more than one area

Table 1 The proportion of interspecific variance in research effort explained by different taxonomic levels (genus, family, order) in sharks

Model	Random effect	% of variance
Model 1	Order	26.98
	Family	23.33
	Genus	4.59
	Residual	45.10
Model 2	Order	21.83
	Family	16.69
	Residual	61.48

Order, family and genus were included as nested random effects where present, and the AIC of the model without any higher taxonomic levels versus the models including just order, family/order, or genus/family/order were compared. Model selection revealed two best models (with $\Delta\text{AIC} < 2$; Model 1, the full model including all three taxonomic levels, with an AIC of 1657.5, and Model 2, including family and order, with an AIC of 1655.6). The percentage of variance explained by each taxonomic level retained in the best models, together with the residual variances, are provided in the table

were the focus of more research, whereas species from the Indian Ocean were less often studied (see Table 3). Intermediate research effort is given to species endemic to the other regions (Table 3).

Biases related to species traits

Except for maximum depth and age at maturity, all tested traits were significantly correlated with research effort (Table 4, see also Fig. 2). Species with a shallow minimum depth, a broader depth range, a broader range of latitudes and a broader habitat breadth had a higher research effort. In addition, body size (both maximum body size or size at birth), maximum longevity, and trophic level were significantly and positively associated with research effort. Species giving birth to free-swimming pups (206 species; 34.2 ± 5.5 papers per species on average) also had a significantly higher research effort than species laying eggs (92 species; 13.2 ± 4.8 papers per species on average; Wilcoxon test: $W = 5938.5$, $p < 0.001$).

In a second set of analyses, I included the predictors of research effort in a single model. Maximum longevity and ages at maturity for each sex were

available for a limited number of species (< 66 , see Table 4), so I excluded these variables from this multivariate analysis to avoid restricting the analysis to a small subset of species. As the Spearman correlation between research effort and maximum depth was non-significant, I did not consider it in this analysis, either. Size at birth and maximum size were highly correlated ($\rho = 0.80$; $p < 0.001$), so size at birth was also excluded to avoid problems with collinearity. Correlations between all other variables were lower than 0.7 (Table S2). Reproductive mode was available for a limited number of species ($n = 298$) and did not significantly predict research effort in the multivariate model ($p = 0.471$) so it was also removed.

I conducted the model selection on the 377 species with data available for all the variables considered in the model (i.e. the fixed effects of minimum depth, depth range, latitudinal range, habitat breadth, maximum size and trophic level and the nested random effects of genus, family, and order). All variables were standardized to a mean of 0 and variance of 1 so that their relative importance could be assessed on a common scale. Geographic region could not be standardized, as it is a categorical variable, so I excluded it from this analysis. Note however that including it in the multiple regression model improved the AIC ($\Delta\text{AIC} = 20.2$, with similar patterns of research effort differences among geographic regions as described above), but that its inclusion did not affect the significance or the relative importance of the other fixed effects. With regards to the random structure, the models including either family nested in order (AIC = 1020.24) or all three taxonomic levels (AIC = 1022.24) had a lower AIC than the null model (AIC = 1027.69) or the model including the order only (AIC = 1025.62). Family nested in order was thus retained as the random structure of the final model (though including genus as well yielded the same final combination of fixed effects). I then used a stepwise procedure to remove non-significant effects one by one, and the final model included three fixed effects, latitudinal range being the most important, followed by body size and finally by minimum depth (see Table 5). Large species occurring in a large range of latitudes, and at shallower minimum depths were more often the focus of scientific articles (Fig. 2). In contrast,

Table 2 Mean research effort per species within each shark order and family

Order	Family	Research effort	SE	N
<i>Carcharhiniformes</i>		18.2	2.9	284
	CARCHARHINIDAE	49.1	10.4	56
	HEMIGALEIDAE	6.1	2	8
	LEPTOCHARIIDAE	5	–	1
	PROSCYLLIIDAE	1.1	0.5	7
	PSEUDOTRIAKIDAE	6.3	4.3	4
	SCYLORHINIDAE	6.7	2.8	153
	SPHYRNIDAE	56	25.9	9
TRIAKIDAE	17.5	4.3	46	
<i>Echinorhiniformes</i>		21	2	2
	ECHINORHINIDAE	21	2	2
<i>Heterodontiformes</i>		17.1	8.9	9
	HETERODONTIDAE	17.1	8.9	9
<i>Hexanchiformes</i>		32.2	12.4	6
	CHLAMYDOSELACHIDAE	30	–	1
	HEXANCHIDAE	32.6	15.2	5
<i>Lamniformes</i>		97.7	28.1	16
	ALOPIIDAE	71	18.2	3
	CETORHINIDAE	215	–	1
	LAMNIDAE	173.6	73.3	5
	MEGACHASMIDAE	64	–	1
	MITSUKURINIDAE	27	–	1
	ODONTASPIDIDAE	38	28.4	4
	PSEUDOCARCHARIIDAE	24	–	1
<i>Orectolobiformes</i>		16.6	6.2	45
	BRACHAELURIDAE	1.5	0.5	2
	GINGLYMOSTOMATIDAE	43.3	38.4	4
	HEMISCYLLIIDAE	13.4	5.2	17
	ORECTOLOBIDAE	6.4	2.5	12
	PARASCYLLIIDAE	1.9	0.4	8
	RHINCODONTIDAE	224	–	1
	STEGOSTOMATIDAE	27	–	1
<i>Pristiophoriformes</i>		2.3	0.6	8
	PRISTIOPHORIDAE	2.3	0.6	8
<i>Squaliformes</i>		14.6	6.3	119
	CENTROPHORIDAE	9.6	3.3	16
	DALATIIDAE	12	4.9	9
	ETMOPTERIDAE	6.6	2.5	45
	OXYNOTIDAE	8.8	4	5
	SOMNIOSIDAE	14.3	4.4	16
	SQUALIDAE	32.2	26.2	28
<i>Squatiniiformes</i>		8.3	1.8	20
	SQUATINIDAE	8.3	1.8	20

SE standard error, N number of species within the taxon

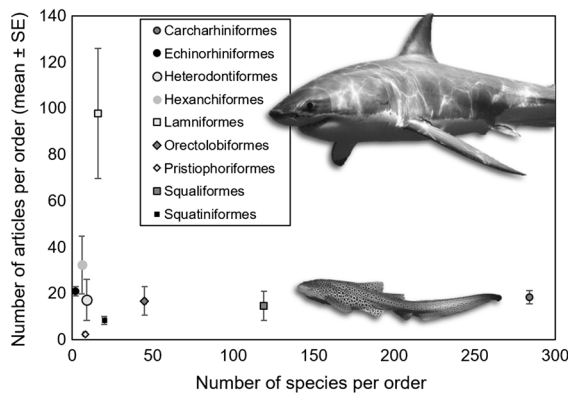


Fig. 1 Mean number of articles per species in relation to the number of species within each order. Order explained about 25% of variance in research effort. Some orders with few species are rarely studied (e.g. Squatiniformes and Pristiophoriformes), despite their phylogenetic uniqueness. For example, the great white shark (*Carcharodon carcharias*, order Lamniformes; top inset), is the third most studied shark species with 412 articles, whereas the leopard shark (*Stegostoma fasciatum*, order Orectolobiformes; bottom inset) is 75th with 27 articles. Photo credit: Jayna DeVore

depth range, trophic level and habitat breadth were not found to be significant predictors of research effort when using the multivariate approach (see

Table 3 Mean research effort per species occurring within each geographic region

Geographic region	Mean research effort	SE	N	Estimate \pm SE	Significance of pairwise comparisons
North-Eastern Atlantic (includes Mediterranean and Black Seas)	63.3	36	11	3.066 \pm 0.404	a
Several	32.4	4.6	251	2.379 \pm 0.209	a
North-Eastern Pacific	16.4	11.3	7	1.726 \pm 0.484	ab
South-Western Atlantic	11	4.9	9	1.227 \pm 0.379	ab
South-Eastern Pacific	6	2.3	13	1.847 \pm 0.436	b
North-Western Atlantic	5.1	3.4	28	0.912 \pm 0.295	b
North-Western Pacific	4	1	58	1.157 \pm 0.249	b
South-Western Pacific	3	0.9	58	0.992 \pm 0.249	b
Western-Indian	1.7	0.5	38	0.738 \pm 0.272	b
Eastern-Indian	1.3	0.1	36	0.799 \pm 0.195	b

Geographic regions are based on the classifications of the International Hydrographic Organization. Species occurring in more than one geographic region were included in the “several” category. Species that are included in a geographic region other than “several” are thus endemic to the region indicated. No species is endemic to the South-Eastern Atlantic. Estimates and their standard errors from the linear model assessing differences in research effort (log transformed) across geographic regions are provided

The letters from the last column illustrate the significance of differences in research effort across geographic regions from Tukey comparisons. Differences in research effort between geographic regions are considered to be significant ($p < 0.05$) if they do not share the same letter (e.g. the category “several” differs from “South-Eastern Pacific” and all other categories below, but not from “North-Eastern Atlantic”, “North-Eastern Pacific” or “South-Western Atlantic”)

SE standard error, N number of species within each region

Table 4 Correlations between research effort and ecological and life history traits in sharks

Correlate of research effort	Spearman rho	p	N
Minimum depth	− 0.552	< 0.001	473
Maximum depth	0.037	0.417	473
Depth range	0.253	< 0.001	473
Habitat breadth	0.326	< 0.001	461
Latitudinal range	0.701	< 0.001	407
Trophic level	0.196	< 0.001	459
Size at birth	0.461	< 0.001	237
Maximum size	0.604	< 0.001	508
Male age at maturity	0.252	0.052	60
Female age at maturity	0.186	0.154	60
Maximum longevity	0.362	0.003	66

N number of species

Significant correlations ($p < 0.05$) are in bold

Table 5). A multimodel inference approach based on the construction of all possible models yielded the same result, as it also revealed one unique best model (with $\Delta AIC < 2$) that included the same variables as the ones included in the best model described above.

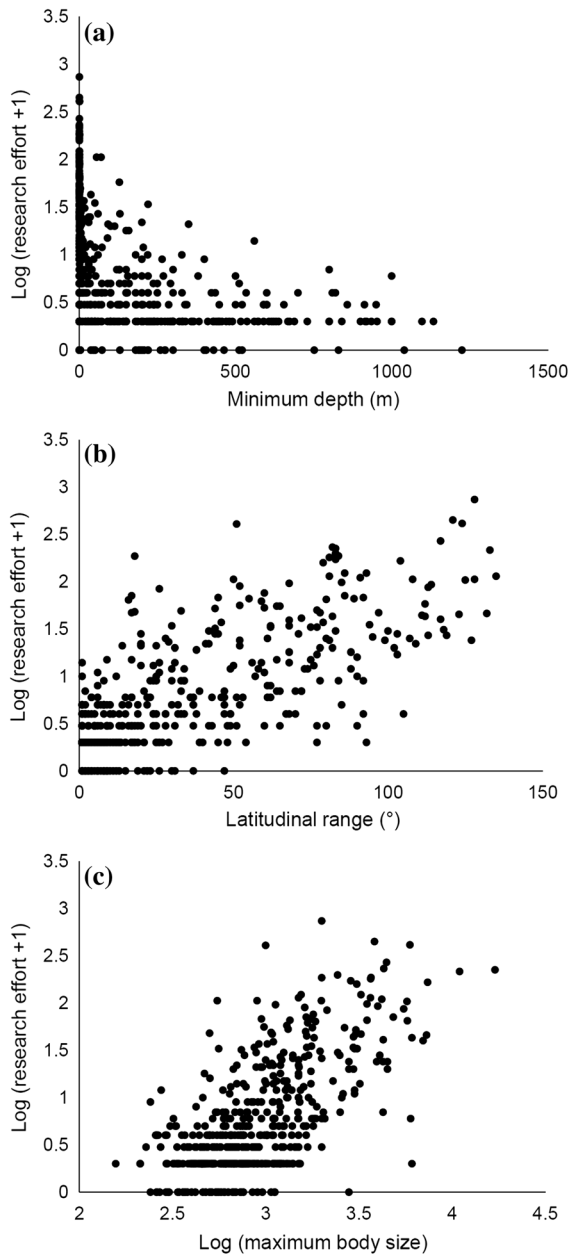


Fig. 2 Associations between research effort (number of articles per species) and **a** minimum depth, **b** latitudinal range and **c** maximum body size. All associations were significant, demonstrating clear, nonrandom biases in shark research effort across the clade

Extinction risk

Unsurprisingly, Data Deficient species and species that have not been assessed by the IUCN had a particularly low research effort (Table 6 and Fig. 3).

After removing these species, I found that research effort significantly varied across extinction risk categories ($F_{4,257} = 6.383$, $p < 0.001$). Species classified as Vulnerable and Near Threatened had a relatively high research effort as compared to Least Concern or Critically Endangered Species, while Endangered species had an intermediate research effort (see Table 6 and Fig. 3). The significance of pairwise comparisons across extinction risk categories is provided in Fig. 3.

Discussion

Research effort varied significantly across the 509 described shark species for which it was estimated (from 0 to over 700 articles per species). This variation was non-random, with important taxonomic and geographic biases, as well as biases associated with a range of ecological and life history traits. Such biases confirm a general trend observed in other taxa (e.g. in birds, Ducatez and Lefebvre 2014), whereby some species and groups of species are targeted for research to the point where they become model organisms, whereas others are neglected. Some of these biases are intentional and scientifically meaningful [e.g. based on the importance of having some well-known model organisms or a focus on specific traits of interest that characterize a few species (Kellogg and Shaffer 1993)], but others are not, and need to be considered before drawing conclusions for an entire taxon. The biases identified in this study and the index provided allow us to better take into account these limitations in future analyses. To illustrate the importance of these biases for future studies, I developed an analysis demonstrating that ignoring the effects of research effort on trait estimations can lead to the interpretation of false positives (type 1 errors) and included this example in the Supplementary Material. To further illustrate how this index should be incorporated into future research I also describe several examples of published comparative studies that included research effort in their analyses (see SM). Briefly, associations of interest to researchers (e.g. hypothetically, the effects of species habitat breadth on parasite diversity) may appear to be significant, even when these effects are actually driven by nonrandom underlying biases in research effort (e.g. habitat generalist species are more well studied, and more well studied species

Table 5 Multivariate model with predictors of research effort in sharks

	Estimate	SE	T	DF	<i>p</i>
Latitudinal range	0.762	0.062	12.200	340	< 0.001
Maximum size	0.346	0.061	5.674	340	< 0.001
Minimum depth	− 0.166	0.056	− 2.964	340	0.003
Depth range	0.112	0.063	1.784	339	0.075
Trophic level	−0.069	0.054	−1.269	339	0.205
Habitat breadth	0.078	0.052	1.496	339	0.136

The parameters and statistics of the best model are given in bold. Depth range, trophic level and habitat breadth were removed from the best model during a stepwise procedure. Each of these three variables was then added separately to the best model to confirm non-significance, providing the parameters and statistics given here for the last three variables. Order and family were included as nested random effects

Table 6 Mean research effort per species within each IUCN red list extinction risk category

Red list status	Mean effort	SE	N	Estimate ± SE
Not assessed	0.8	0.1	37	–
Data deficient	6.1	1	210	–
Critically endangered	8.5	2.5	10	1.998 ± 0.461
Endangered	26.8	16.2	14	2.002 ± 0.603
Vulnerable	68.8	19.1	47	2.850 ± 0.507
Near threatened	39.6	8.4	66	2.708 ± 0.494
Least concern	17.4	4	125	1.831 ± 0.479

Not assessed = species not included in the IUCN red list

Estimates and their standard errors from the linear models assessing differences in research effort (log transformed) across extinction risk categories are provided (note that the model excluded data deficient and not assessed species) (see also Fig. 3)

SE standard error, N number of species within each category

have more documented parasites, so generalist species appear to have more parasites). The inclusion of research effort in such analyses is therefore critical in order to avoid drawing erroneous conclusions about the significance of associations of interest; the analysis in the Supplementary Material uses a database on parasite prevalence to demonstrate that the inclusion of research effort as a confounding variable can indeed drastically change the conclusions reached via comparative analyses.

Geographic region and accessibility to researchers were major predictors of interspecific variation in research effort (Tables 3, 4, 5). Geographic variation in research effort was predominantly predicted by differences in research investment across continents: the North Eastern Atlantic (including the

Mediterranean Sea) and North Eastern Pacific zones, adjacent to Europe and North America, were the most studied areas. In contrast, the Indian Ocean regions were much less studied, despite the high vulnerability of species and the importance of fishing pressures from the shark fin industry in these areas (Dulvy et al. 2014). The lack of overlap between areas of high conservation importance and research effort is already known in other taxa (Brito and Oprea 2009; de Lima et al. 2011; Ducatez and Lefebvre 2014) but is particularly concerning in sharks, as the main threat responsible for their decline is (over)exploitation. Overcoming this threat requires detailed knowledge of the response of each population to exploitation so that effective protection measures can be designed, underlining the need for improving our knowledge of

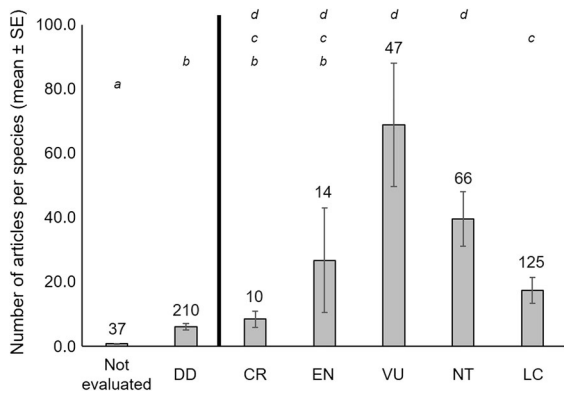


Fig. 3 Mean research effort per species within each extinction risk category. The black line segregates species which extinction risk is not known (not evaluated and *DD* data deficient) from species which extinction risk is known (*CR* critically endangered, *EN* endangered, *VU* vulnerable, *NT* near threatened, *LC* least concern). The number of species is provided for each category. The letters a to d are used to illustrate the significant differences in research effort across extinction risk categories from Tukey comparisons. Categories are considered to be significantly different ($p < 0.05$) if they do not share the same letter (e.g. *LC* differs from *NT*, *VU*, *DD* and not evaluated but not from *CR* or *EN*)

species that occur in areas with high fishing pressures. One surprising geographic pattern is the low research effort of species endemic to the North Western Atlantic zone (5.1 papers per species for the 28 species endemic to that area, Table 3). A likely explanation is that most of these species are catsharks and lantern sharks (see Table S1), deep-sea species that are difficult to study.

Latitudinal range, habitat breadth and depth range were all correlated with research effort, likely because species occurring in a larger range of latitudes, habitats and depths are more often encountered by researchers, and thus easier to study. The association between geographic range size and year of discovery, with species that were discovered earlier having larger ranges (Randhawa et al. 2015), is in line with these results (see also Garamszegi and Møller 2012; Ducatez and Lefebvre 2014 for similar patterns in birds). Similarly, the finding that research effort decreased with minimum depth is very likely attributable to the logistic and financial challenges of studying deep-sea species. The latter result is also in line with the study by Randhawa et al. (2015), which demonstrated that shark species occurring in shallow waters are described earlier than species from deeper

waters. Minimum depth and latitudinal range were also two of the only three parameters retained in the multivariate analysis, underlining the central importance of parameters that are related to the ease of study in shaping shark research effort.

The third parameter included in the multivariate analysis was maximum body size, larger species being more frequently investigated. This focus on large species is likely to have a few explanations: first, charismatic species, which often attract larger research budgets due to increased public interest (Colléony et al. 2017; Albert et al. 2018), also tend to be large species, such as the great white shark, the tiger shark or the whale shark. Second, larger species may be easier to study. Until recently, only large species could be equipped with transmitters, a potentially important source of research bias in the study of shark habitat use and movement ecology. Direct observations and individual recognition are also easier on larger species. Third, large sharks are likely more profitable and particularly targeted by fisheries, increasing the importance of studying these species. Fourth, large shark species are involved in another human activity, tourism. The development of tourist shark diving activities requires a good understanding of the observed species, and favors observations that can sometimes be used by the scientific community (Mieras et al. 2017). Finally, the handful of shark species involved in fatal attacks on humans are large species (ISAF 2019), the prevention of such attacks being an important source of funding for research on these species. This targeting of large species indirectly introduces important biases towards other traits, including life history traits such as longevity and size at birth, and ecological traits such as the trophic level. In addition, these differences in research effort can affect some trait values attributed to the species. Estimates of maximum body size, maximum longevity, latitudinal range, habitat breadth and depth range are, among others, likely to increase as the amount of research on a given species increases, as both maximum values and ranges are affected by the sampling effort. The correlations between these variables and research effort may thus also be related to biases in parameter estimates resulting from differences in research effort (Garamszegi and Møller 2010, 2011, 2012).

As mentioned above, research efforts are particularly low in some areas of high conservation

importance. In addition, species risk of extinction was associated with research effort, though not linearly: species with an intermediate risk of extinction (classified as Vulnerable or Near Threatened) had a particularly high research effort, whereas species that are considered as either Critically Endangered or Least Concern were rarely studied (Fig. 3). The distribution of research effort according to species extinction risk appears better balanced in sharks than in birds, where the research effort of Least Concern species was twice as high as that for species from any other extinction risk category (Ducatez and Lefebvre 2014). The very low effort assigned to Critically Endangered and Endangered species in sharks is concerning, however, though it is likely explained by the difficulties involved in studying populations on the verge of extinction. It is also notable that the main traits predicting variation in research effort are also major predictors of variation in extinction risk in sharks (Dulvy et al. 2014). Minimum depth and body size are two key factors that affect both the number of studies reported per species and the risk of extinction (Dulvy et al. 2014). The similarity between fisheries and research targets is likely responsible for these associations, with both targeting larger species and having limited access to species occurring at deeper depths. As a result, traits that attract research also tend to attract overfishing, creating an association between extinction risk and research effort.

This paper aims to highlight the non-random distribution of research effort across shark species, while assessing the nature and importance of research effort biases. It provides research effort estimates for 509 extant shark species and shows that, in addition to taxonomy and geographic distribution, latitudinal range, maximum body size and minimum depth are the main predictors of research effort. These biases need to be considered in reviews, meta-analyses and comparative analyses as they may affect the general conclusions by skewing our understanding towards species with characteristics that attract more research. As fisheries are switching towards the deep-sea, there is an urgent need to develop research programs aimed at assessing how populations of these poorly known species will respond to exploitation (Morato et al. 2006; Robison 2009). The increase of deep-sea exploration and exploitation abilities thus represents both a great opportunity and a major challenge for shark biologists. If these research effort biases are

neglected, relatively unknown species that need better protection may suffer from a lack of informed conservation policies.

Acknowledgements I thank Jayna DeVore for providing photos, for her help designing Fig. 1, and for her very helpful comments and language editing of the manuscript. I also thank two anonymous reviewers for suggestions that helped in improving the manuscript.

Compliance with ethical standards

Conflict of interest The author declares that no competing interests exist.

References

- Albert C, Luque GM, Courchamp F (2018) The twenty most charismatic species. *PLoS ONE* 13:e0199149. <https://doi.org/10.1371/journal.pone.0199149>
- Arnold C, Nunn CL (2010) Phylogenetic targeting of research effort in evolutionary biology. *Am Nat* 176:601–612
- Barker MJ, Schluessel V (2005) Managing global shark fisheries: suggestions for prioritizing management strategies. *Aquat Conserv Mar Freshw Ecosyst* 15:325–347. <https://doi.org/10.1002/aqc.660>
- Brito D, Oprea M (2009) Mismatch of research effort and threat in avian conservation biology. *Trop Conserv Sci* 2:353–362
- Colléony A, Clayton S, Couvet D et al (2017) Human preferences for species conservation: animal charisma trumps endangered status. *Biol Conserv* 206:263–269. <https://doi.org/10.1016/j.biocon.2016.11.035>
- Compagno L, Dando M, Fowler S (2005) *Sharks of the world*. Princeton University Press, Princeton
- de Lima RF, Bird JP, Barlow J (2011) Research effort allocation and the conservation of restricted-range island bird species. *Biol Conserv* 144:627–632. <https://doi.org/10.1016/j.biocon.2010.10.021>
- Ducatez S, Lefebvre L (2014) Patterns of research effort in birds. *PLoS ONE* 9:e89955. <https://doi.org/10.1371/journal.pone.0089955>
- Ducatez S, Shine R (2017) Drivers of extinction risk in terrestrial vertebrates. *Conserv Lett* 10:186–194. <https://doi.org/10.1111/conl.12258>
- Dulvy NK, Fowler SL, Musick JA et al (2014) Extinction risk and conservation of the world's sharks and rays. *Elife* 3:e00590
- Ferretti F, Worm B, Britten GL et al (2010) Patterns and ecosystem consequences of shark declines in the ocean. *Ecol Lett* 13:1055–1071. <https://doi.org/10.1111/j.1461-0248.2010.01489.x>
- Froese E, Pauly D (2017) *FishBase*. <https://www.fishbase.org>. Accessed Jan 2017
- Garamszegi LZ, Møller AP (2010) Effects of sample size and intraspecific variation in phylogenetic comparative studies:

- a meta-analytic review. *Biol Rev* 85:797–805. <https://doi.org/10.1111/j.1469-185X.2010.00126.x>
- Garamszegi LZ, Møller AP (2011) Nonrandom variation in within-species sample size and missing data in phylogenetic comparative studies. *Syst Biol* 60:876–880. <https://doi.org/10.1093/sysbio/syr060>
- Garamszegi LZ, Møller AP (2012) Untested assumptions about within-species sample size and missing data in interspecific studies. *Behav Ecol Sociobiol* 66:1363–1373. <https://doi.org/10.1007/s00265-012-1370-z>
- González-Suárez M, Lucas PM, Revilla E (2012) Biases in comparative analyses of extinction risk: mind the gap. *J Anim Ecol* 81:1211–1222. <https://doi.org/10.1111/j.1365-2656.2012.01999.x>
- Heithaus MR, Frid A, Wirsing AJ, Worm B (2008) Predicting ecological consequences of marine top predator declines. *Trends Ecol Evol* 23:202–210. <https://doi.org/10.1016/j.tree.2008.01.003>
- IHO (1953) Limits of oceans and seas, 3rd. International Hydrographic Organization, Monte-Carlo
- ISAF (2019) International shark attack file. <https://www.floridamuseum.ufl.edu/shark-attacks/>. Accessed Feb 2019
- IUCN (2017) The IUCN red list of threatened species. Version 2017-3. <http://www.iucnredlist.org>. Accessed Jan 2017
- Kellogg EA, Shaffer HB (1993) Model organisms in evolutionary studies. *Syst Biol* 42:409–414. <https://doi.org/10.1093/sysbio/42.4.409>
- Mieras PA, Harvey-Clark C, Bear M et al (2017) Chapter five—the economy of shark conservation in the northeast pacific: the role of ecotourism and citizen science. In: Larson SE, Lowry D (eds) *Advances in marine biology*. Academic Press, Cambridge, pp 121–153
- Molina JM, Cooke SJ (2012) Trends in shark bycatch research: current status and research needs. *Rev Fish Biol Fish* 22:719–737. <https://doi.org/10.1007/s11160-012-9269-3>
- Morato T, Watson R, Pitcher TJ, Pauly D (2006) Fishing down the deep. *Fish Fish* 7:24–34
- Randhawa HS, Poulin R, Krkošek M (2015) Increasing rate of species discovery in sharks coincides with sharp population declines: implications for biodiversity. *Ecography* 38:96–107. <https://doi.org/10.1111/ecog.00793>
- Rasmussen T, Randhawa H (2018) Host diet influences parasite diversity: a case study looking at tapeworm diversity among sharks. *Mar Ecol Prog Ser* 605:1–16. <https://doi.org/10.3354/meps12751>
- Robison BH (2009) Conservation of deep pelagic biodiversity. *Conserv Biol* 23:847–858. <https://doi.org/10.1111/j.1523-1739.2009.01219.x>
- Roff G, Doropoulos C, Rogers A et al (2016) The ecological role of sharks on coral reefs. *Trends Ecol Evol* 31:395–407. <https://doi.org/10.1016/j.tree.2016.02.014>
- Simpfendorfer CA, Kyne PM (2009) Limited potential to recover from overfishing raises concerns for deep-sea sharks, rays and chimaeras. *Environ Conserv* 36:97. <https://doi.org/10.1017/S0376892909990191>
- Smith SE, Au DW, Show C (1998) Intrinsic rebound potentials of 26 species of Pacific sharks. *Mar Freshw Res* 49:663. <https://doi.org/10.1071/MF97135>
- Sol D, Duncan RP, Blackburn TM et al (2005) Big brains, enhanced cognition, and response of birds to novel environments. *Proc Natl Acad Sci USA* 102:5460–5465
- Weigmann S (2016) Annotated checklist of the living sharks, batoids and chimaeras (Chondrichthyes) of the world, with a focus on biogeographical diversity: annotated global checklist of chondrichthyes. *J Fish Biol* 88:837–1037. <https://doi.org/10.1111/jfb.12874>
- Weigmann S (2017) Reply to Borsa (2017): Comment on ‘Annotated checklist of the living sharks, batoids and chimaeras (Chondrichthyes) of the world, with a focus on biogeographical diversity by Weigmann (2016)’. *J Fish Biol* 90:1176–1181. <https://doi.org/10.1111/jfb.13234>
- Worm B, Davis B, Kettner L et al (2013) Global catches, exploitation rates, and rebuilding options for sharks. *Mar Policy* 40:194–204. <https://doi.org/10.1016/j.marpol.2012.12.034>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.