Vol. 280: 285–293, 2004 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published October 14

Modelling post-release mortality of loggerhead sea turtles exposed to the Hawaii-based pelagic longline fishery

Milani Chaloupka[1,*], **Denise Parker**[2], **George Balazs**[3]

[1]Ecological Modelling Services Pty Ltd, PO Box 6150, University of Queensland, St Lucia, Queensland 4067, Australia
[2]Joint Institute for Marine and Atmospheric Research, 8604 La Jolla Shores Drive, La Jolla, California 92037-0271, USA
[3]National Marine Fisheries Service, Pacific Islands Fisheries Science Center, 2570 Dole Street, Honolulu, Hawaii 96822-2396, USA

ABSTRACT: Loggerhead sea turtles *Caretta caretta* are an endangered species exposed to anthropogenic hazards such as pelagic longline fisheries. Many loggerheads caught in these fisheries are alive when released from the gear, but many probably die soon after because of hook injuries or line entanglement. Robust estimates of post-release mortality are essential for stock assessment and evaluating the benefit of releasing turtles caught alive in the gear, yet none are available for any sea turtle species. Here, the post-release mortality of 40 loggerheads caught in the Hawaii-based pelagic longline fishery was investigated using satellite telemetry deployed by a National Marine Fisheries Service (NMFS) observer program. We modelled time-to-failure of all transmitters using nonparametric statistical modelling (Kaplan-Meier-Turnbull, local regression) to derive survival and hazard functions for light and deep hooked loggerheads. There was a significant difference between the survival functions for light and deep hooked loggerheads within 90 d of release, but no difference between survival functions after this time. But satellite transmitters fail for many reasons (defects, battery failure, transmitter detachment, turtle death), which results in a hazard function that confounds these competing risks. Hence we propose that it might not be possible to infer true post-release mortality based on satellite telemetry unless the cause of each transmitter failure is known, which is rarely the case. We discuss other survey design and statistical modelling challenges involved in the evaluation of post-release mortality based on satellite telemetry.

KEY WORDS: Loggerhead sea turtles · Pelagic longline fisheries · Satellite telemetry · Post-release mortality · Failure time modelling · Competing risks

INTRODUCTION

There is increasing global concern about the incidental mortality of sea turtles caught in commercial fisheries (Hall et al. 2000). Some populations of loggerhead turtle *Caretta caretta* are in serious decline, especially in Pacific waters where there are 2 distinct genetic stocks (see Fig. 1), an Australian stock and a Japanese stock (Bowen et al. 1994). The decline of both stocks has been attributed to several hazards, including fox predation of eggs, nesting habitat destruction and exposure to coastal and pelagic fisheries (Chaloupka 2003, Kamezaki et al. 2003) and perhaps direct harvesting of the Japanese stock (Gardner & Nichols 2001).

Most studies of fishery-related loggerhead mortality have focused on estimates of pre-release mortality (Poiner & Harris 1996, Cheng & Chen 1997, Julian & Beeson 1998, Laurent et al. 1998, Slater et al. 1998, McCracken 2000), which is usually attributable to drowning (Work & Balazs 2002). Many sea turtles caught in fishing gear are alive when released from

---

*Email: m.chaloupka@uq.edu.au

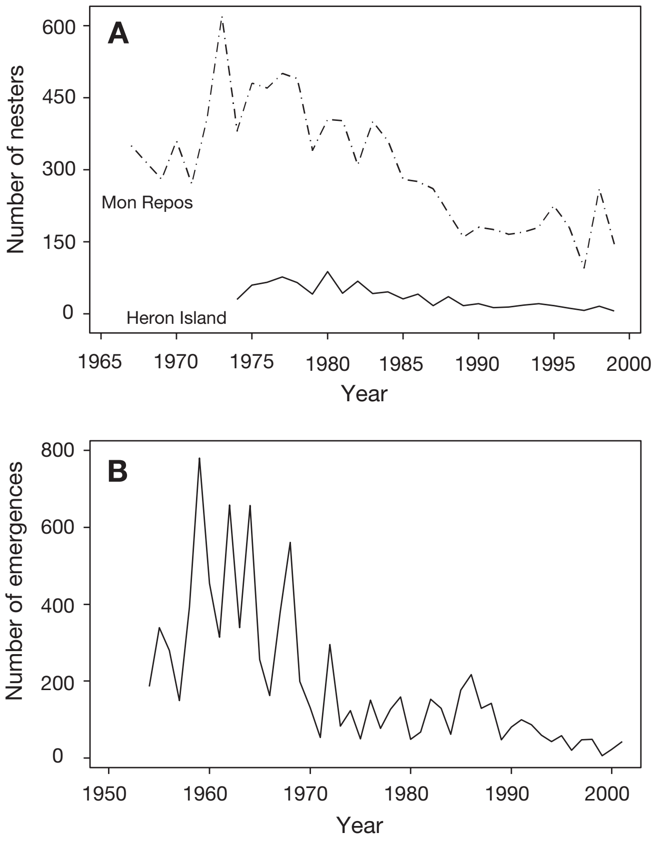

Fig. 1. *Caretta caretta*. Long-term trends in the female nesting abundance of the 2 Pacific loggerhead stocks. (A) Number of female nesters recorded for Australian stock loggerheads at the Heron Island (Chaloupka & Limpus 2001) and Mon Repos rookeries (Limpus & Limpus 2003). (B) Number of female beach emergences or haulouts recorded for Japanese stock loggerheads at the Kamouda rookery in the Tokushima Prefecture (Kamezaki et al. 2003)

the gear, but it is assumed that many will die anyway because of injuries caused by the hooks or line entanglement (Aguilar et al. 1995, Hall et al. 2000).

However, very few studies have empirically addressed the issue of post-release mortality for sea turtles by monitoring the behavior and short-term survival of sea turtles released alive from the fishing gear (Parker et al. in press). Reliable information on post-release mortality could be useful for loggerhead stock assessment and for evaluating the benefit of releasing loggerheads caught alive in the longline gear (Chaloupka & Limpus 2002, Chaloupka 2003).

Satellite telemetry has been proposed as a useful technology for evaluating post-release mortality of sea turtles (Balazs & Pooley 1994, Bjorndal et al. 1999, Parker et al. in press) and for other large and mobile pelagic species such as billfish (Goodyear 2002, Graves et al. 2002). Satellite tracking of individual loggerheads in North Pacific waters has already proved useful for investigating migratory pathways of mature turtles (Sakamoto et al. 1997) and the post-release dispersal and pelagic ecology of immatures caught in longline gear (Polovina et al. 2003b).

In the present study satellite telemetry was used to investigate the post-release mortality of loggerheads caught in the Hawaii-based pelagic longline fishery that operated in the central North Pacific between 1997 and 2000. All loggerheads caught in this fishery were from the Japanese genetic stock, as there is no evidence of any Australian loggerheads (Dutton et al. 1997). More details on the Hawaii-based longline fishery can be found in Polovina et al. (2003a).

MATERIALS AND METHODS

Satellite transmitter deployment. Trained NMFS observers were randomly assigned aboard ca. 5% of the Hawaii-based commercial longline fishing fleet from 1997 to 2000 (Parker et al. in press). The observers retrieved 267 sea turtles captured on longline gear, 168 of which were loggerheads and of this number, 2 were dead on retrieval from the gear. For each trip, each observer was supplied with 1 or 2 satellite transmitters to attach to any hard-shelled sea turtle retrieved alive during their 2 to 6 wk observation period.

All turtles brought on board alive were scored by the NMFS observer as either (1) deep hooked, if the longline hook was present in the gastrointestinal tract caudal to the glottis, or (2) light hooked, if the hook was lodged in the mouth or externally, or the turtle was entangled in the line (see Work & Balazs 2002). If the loggerhead was light hooked, then the study protocol required the hook to be removed. If deep hooked then the hook was left in place in accordance with veterinarian recommendations (Balazs et al. 1995).

The observers attached Argos-linked satellite transmitters to 40 loggerheads that were released alive after incidental capture by the longline gear (Parker et al. in press). Transmitter assignment was based on the following criteria given transmitter availability: (1) the turtle was alive when retrieved, (2) the shell size was big enough for transmitter attachment (>40 cm straight carapace length, SCL), and (3) the sea conditions and weather (wind, rain) were suitable to allow attachment using fiberglass strips and polyester resin as described by Balazs et al. (1996).

A total of 38 loggerheads were fitted with Telonics ST-10 or ST-18 transmitters, while a further 2 were fitted with Wildlife Computers satellite-linked depth recorder (SDR-T10) transmitters (see Polovina et al. 2003a for more details). Each loggerhead with a transmitter was then released alive by the observers shortly

Table 1. Summary of the 40 transmitters deployed on the 40 loggerhead sea turtles released alive in the study and arranged by hooking position and transmitter duty cycle

Hook position	Transmitter duty cycle (on/off, in h) 2/4	12/48	SDR	24/216	Total
Light	9	1	1	2	13
Deep	12	5	1	9	27

after attachment with no apparent signs of morbidity. In this study, 4 satellite transmitter duty cycles were used: (1) 2/4 h, (2) 12/48 h, (3) 24/216 h, and (4) the SDR transmitters. A 24/216 duty cycle means that transmission to the satellite network comprised a cycle of 24 h on followed by 216 h (or 9 d) off. The longer duty cycles were used in an attempt to conserve battery power and extend the tracking period.

A summary of the 40 satellite transmitters arranged by hooking position and transmitter duty cycle is given in Table 1. The data set for all 40 transmitters is available on request from G. H. Balazs (NMFS, Honolulu, Hawaii). More details on the NMFS observer program and deployment of these transmitters are given in Parker et al. (in press) while Polovina et al. (2003b) analyzed post-release movement patterns of some of these loggerheads using satellite telemetry and remote sensing data. Note that Polovina et al. (2003b) did not address post-release mortality of any of the satellite-tracked loggerheads in their study.

Statistical modelling approach. The satellite transmission duration in days from release until all transmissions ceased was determined for each of the 40 loggerheads summarized in Table 1. The transmission duration only reflects the time period until failure of the transmitter and is thus not a direct measure of the short-term survival duration of the tag-and-released turtle. This is because satellite transmissions can cease for many reasons, such as transmitter loss due to improper attachment, transmitter defects, battery failures, and of course death of the turtle carrying the transmitter. Post-release loggerhead death could be due to injuries sustained during capture in the longline gear or due to natural mortality.

There are many competing risks that could lead to transmitter failure so identifying cause-specific transmission failure is extremely difficult, let alone inferring cause-specific loggerhead mortality. Chaloupka & Musick (1997) discuss in more detail these sorts of competing risks in the context of tag loss from sea turtle capture-mark-recapture studies while Goodyear (2002) discusses similar issues in relation to billfish tag-and-release studies.

Not only do these data confound cause-specific risks, there are also significant data censoring issues involved such as left and right censoring and interval censored data structures. Right (or left) censored survival time data occur when there are subjects in the sample for which only an upper (or lower) bound on the survival (transmitter failure time) time is known. Right censored data usually occur because the study ends before all subjects in the sample fail or die. Left censored data can occur when some subjects fail soon after the study starts, but before the first time that an event such as death (transmitter failure) can be recorded. Interval censored data occur when there are subjects in the sample for which it was only possible to record a lower and an upper bound (an interval) on the survival time (transmitter failure time). These 3 data structures can be analysed in more detail as follows:

Right censoring: It is important to note that not all the data reflect an event time, which is the time period in days until the known transmission failure. The original data set used by Parker et al. (in press) comprised 39 of the 40 transmitters used here, but 4 were excluded as they were still transmitting at the time of the analysis (~experiment-end). Excluding these data is unnecessary and could result in biased survival function estimates. The data are easily included by identifying the transmission duration times for those 4 transmitters as right censored. In other words, the time period until failure would be at least as long as the time estimated at the end of the experiment. Preliminary analyses of the data used right censoring of these 4 turtles but now all 40 transmitters have ceased functioning for various reasons. Therefore right censoring is not needed here, but is a potential feature of the analysis undertaken in this study.

Left censoring: It is important to note that 10 of the 40 transmitters failed to produce any transmissions and so were identified as left censored observations here to reflect that failure occurred, for whatever reason, sometime between the time of release and the next available time the transmitter could provide a transmission to the Argos network. This is particularly important if failure occurred for a transmitter that was programmed on a long cycle such as 24/216, since the failure could have occurred within the first 24 h after release or perhaps for as long as 9 d after release.

We adopted a left censored scoring to reflect that failure occurred before the first transmission to the Argos network was possible (see Meeker & Escobar 1998 for more details on left censored mechanisms). Parker et al. (in press) discarded these early failures, but we find this unnecessary and furthermore, it could lead to biases in the survival functions and a misunderstanding of the underlying hazard or time-specific mortality function, as it is also possible that the turtle did in fact die soon after release from the longline gear. All we can determine is that there were a significant number of early failures between release (time = 0) and first successful transmission, but the cause for each failure was unknown.

Interval censored data: Not only could the data set comprise both right and left censored data—it also comprised variable interval data, which is a significant complication. Left censoring is also a form of interval censoring but with the lower bound being zero. Recall that the transmission data set comprised 2 SDR transmitters and 38 transmitters programmed with either a 2/4, 12/48 or 24/216 duty cycle. The longer duty cycles (12/48, 24/216) indicated that it was not possible to acquire daily fixes and that a transmission signal was only possible within a time window perhaps as long as 10 d apart. Hence transmitter durations derived from the longer duty cycle satellite transmitters represent interval censored data and few survival models exist (nonparametric, semiparametric or parametric) that are capable of dealing with such data.

Therefore, this data set of 40 transmission failure times for 27 deep hooked and 13 light hooked loggerheads now comprises left and interval censored data, and originally also comprised right censored data. Lawless (1982) provides a thorough discussion of censored data types and various approaches for dealing with such data in a statistical modelling framework.

We modelled the distribution of the transmitter failure times for the deep and light hooked groups using the extended and generalized form of the non-parametric Kaplan-Meier survival function estimator developed by Turnbull (1976). The standard form of the Kaplan-Meier estimator accounts only for right censored data (see review in Lawless 1982). The generalized Kaplan-Meier-Turnbull (KMT) estimator, also known as the Peto-Turnbull estimator (Meeker & Escobar 1998), enabled us to derive group-specific survival functions from the failure times without assumptions about the form of the underlying hazard function while accounting for left, right and interval censored data.

The survival function here is the probability of the Argos network, for any specified time *t* since post-release, recording a signal from the transmitter at least to time *t*. All KMT survival functions were estimated using the kaplanMeier(censor(...)) function in SPLUS (MathSoft 1999), which is an extension of the **S** statistical language. Complementary log-log confidence bands were used to constrain the KMT survival function estimates to the [0,1] interval. We then used the KMT estimated group-specific transmitter survival functions to infer group-specific post-release survival for this sample of loggerheads.

It was not possible to use the standard G^p family of tests (Harrington & Fleming 1982) of the difference between the KMT survival functions because, as we shall see, the hazard functions are neither proportional nor monotonic, but most probably of the bathtub type (Glaser 1980). A bathtub hazard function is a function that refers to a broad class of U- or J-shaped hazard functions and is the term widely applied to such functions in the medical, reliability, engineering, actuarial, economic and ecological literature (Glaser 1980, Lawless 1982, Paranjpe & Rajarshi 1986, Meeker & Escobar 1998). Instead, we used visual examination of the KMT survival functions and 95% complementary log-log confidence bands to evaluate any group-specific survival function differences.

The hazard function gives for any specified time *t* the instantaneous risk of failure at time *t* among transmitters still operating at least until time *t*. We used local log-quadratic likelihood regression smoothing (Loader 1999) of the transmitter failure times to derive preliminary estimates of the underlying group-specific hazard rate functional form using the locfit() library in **R** (Ihaka & Gentleman 1996). The hazard function reflects the time-specific mortality or failure rate of the transmitters. This local regression smoothing approach cannot fully account for the interval and censored data structure, so it is only useful for exploring the functional form rather than deriving robust estimates of the hazard function. Moreover, while this is the largest data set of this type, it is nonetheless far too small for attempting to derive robust estimates of the hazard function.

In deriving the KMT survival functions it was assumed that the censoring mechanism (left, right, interval) is noninformative and not due to some characteristic of a subset of transmitters or loggerheads (see Lawless 1982 for a discussion of this important issue). This assumption can be tested using a multinomial logit regression with the vector that identifies the transmitter censoring type (0 = right, 1 = actual event, 2 = left and 3 = interval) as the response variable to be conditioned on covariates such as duty-cycle, loggerhead size and tagging cohort year. Unfortunately, the data set is too small for robust evaluation of noninformative censoring.

RESULTS

All 40 loggerheads were probably immature and ranged in carapace size from 41 to 83 cm SCL (median size = 58 cm SCL, interquartile range [see Cleveland 1993] = 53 to 65 cm). All turtles were smaller than the known adult size for this stock (Hatase et al. 2002). The duration of satellite transmissions for the 40 loggerheads ranged from 0 to 597 d (median duration = 97 d, interquartile range 3 to –154 d). Distance traveled ranged from 0 to 13 864 km, with a mean distance of ca.

1311 km (see Parker et al. in press, and Polovina et al. 2003b for details on estimating location and distance travelled).

Recall that 27 of the 40 loggerheads were recorded as deep hooked, while 13 were recorded as light hooked. The expected KMT survival functions for these 2 groups given the left and interval censoring are shown in Fig. 2A with 95% complementary log-log confidence bands, and then again in Fig. 2B without the confidence bands to avoid visual clutter. There is little or no overlap between the 2 survival functions (Fig. 2A) between ca. 60 and 90 d of release from the longline gear. The confidence bands are necessarily broad because of the small sample size, but are suggestive of a meaningful difference between the 2 survival curves until around 90 to 100 d post-release (Fig. 2A).

The general behavior of the 2 functions is clearer in Fig. 2B, where the deep hooked survival function declines immediately following release. The light hooked survival function declines after a short delay and then most rapidly after 90 d post-release. Median transmitter failure time for the deep hooked group was ca. 50 d post-release, while median transmitter failure time for the light hooked group was ca. 100 d post-release (Fig. 2B). The group-specific survival functions converge by ca. 120 d post-release.

The survival functions for the 2 groups suggest very different underlying hazard functions (Fig. 2C), which display distinct bathtub shapes (Glaser 1980, Paranjpe & Rajarshi 1986) with a period of early failure followed by a period of low mortality or failure and then followed by an accelerating period of failure after at least a couple of months. The underlying bathtub-type hazard functions (U- or J-shaped functions) and the fact that they cross each other (Fig. 2C) are the reasons why it was not possible to use the G^p family of rank tests (Harrington & Fleming 1982) to test for a difference between the KMT survival functions in Fig. 2A.

These hazard functions were not well defined because of the small sample size and because only the deep hooked turtles had transmitters functioning >200 d. The longer-term hazard function for the deep hooked group is due to 3 loggerheads released towards the end of the study, all with 12/48 duty cycle transmitters attached. Therefore, little attention should be given to the longer-term behavior of the deep hooked function or to the various intermediate bumps in either function, except for the distinct and temporary rise in the hazard or time-specific failure rate for the deep hooked group around 50 d since release (Fig. 2C). Whether this reflects an abrupt increase in time-specific mortality for the deep hooked loggerheads or transmitter failure, due perhaps to battery failure, is unknown.

Recall that the study comprised the satellite transmission durations for 40 transmitters deployed over 4 annual sampling cohorts and using 4 different transmission duty cycles. A total of 21 transmitters were

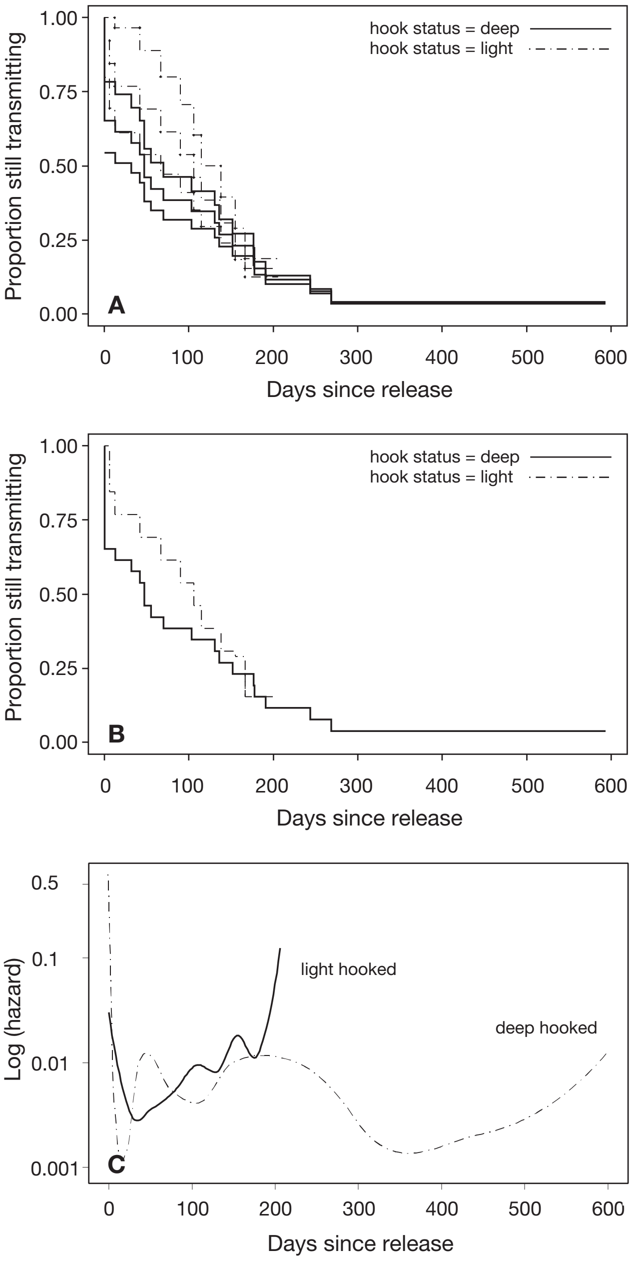

Fig. 2. *Caretta caretta.* Kaplan-Meier-Turnbull (KMT) survival functions or smoothed hazard functions for the 40 satellite tracked deep and light hooked loggerheads. (A) Expected group-specific survival functions including 95% complementary log-log confidence intervals. (B) Expected survival functions only without confidence intervals. (C) Expected local log-quadratic likelihood hazard functions for the deep and light hooked loggerhead groups derived using the locfit() library (see Loader 1999)

programmed with a 2/4 duty cycle and there were 2 SDR transmitters; 17 transmitters were programmed with longer duty cycles comprising 11 on 24/216 and 6 on 12/48. Furthermore, 9 transmitters were deployed in 1997, 16 in 1998, 7 in 1999, and 8 in 2000. The satellite transmitters programmed with the longer duty cycles were deployed mainly toward the end of the study, resulting in potential duty cycle-sampling cohort confounding. It would be preferable to condition the group-specific survival functions on duty cycle and sampling cohort effects, but this is not feasible given the very small sample size.

Nonetheless, we can explore the main effects of duty cycle and sampling cohort on the survival functions for the 40 satellite transmission durations. Fig. 3A shows the duty cycle-specific survival functions with 95% complementary log-log confidence bands included for the 2/4 and 24/216 duty cycles that had reasonable sample sizes. It would seem that duty cycle has a significant effect on the survival functions with the longest cycle (24/216: median failure time ca. 30 d) having significantly lower survival than the shortest cycle (2/4: median failure time ca. 60 d).

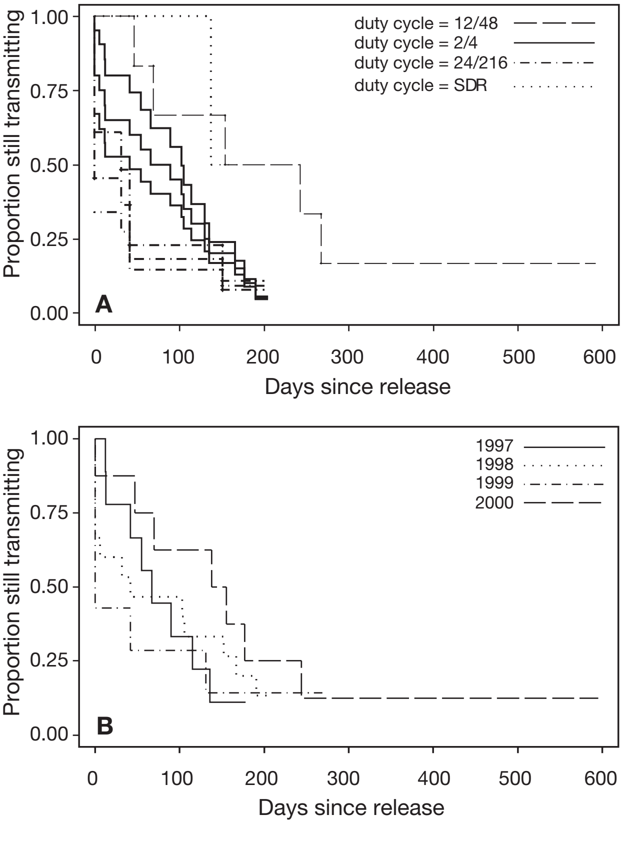

Fig. 3. *Caretta caretta*. Kaplan-Meier-Turnbull (KMT) survival functions for the 40 satellite tracked deep and light hooked loggerheads. (A) Expected duty cycle-specific survival functions for the 4 programmed duty cycles, including 95% complementary log-log confidence intervals for the 2/4 and 24/216 cycles. Confidence intervals are not included for 12/48 or SDR cycles due to very small sample sizes and to avoid visual clutter. (B) Expected year-specific survival functions for the 4 annual sampling cohorts without confidence intervals

However, as already mentioned, the duty cycle and sampling cohort effects are confounded. Fig. 3B shows the sampling cohort-specific survival functions without confidence bands to avoid visual clutter. The 1999 sampling cohort is the lowest survival function simply because there were no light hooked loggerheads in that year, and there were a disproportionate number of both 24/216 duty cycle transmitters deployed and left censored transmission durations recorded. Recall that the survival function modelling here assumes that the censoring mechanisms are noninformative, but it would seem that during 1999 this was probably not a valid assumption. This might also be so for the 2000 sampling cohort when there was a disproportionate number of the interval censored durations, although there was no apparent bias with respect to censor type or whether the loggerhead was deep or light hooked. Again, the sample size in this study is too small to draw strong inferences about differences between transmitter failure times.

DISCUSSION

The satellite tracking of 40 loggerheads released alive in the Hawaii-based longline fishery suggests that there is a difference between the survival functions of transmitters attached to deep and light hooked turtles. The difference in the survival functions occurs within ca. 90 d of release, with no apparent difference between survival functions after this time (Fig. 2B). However, the survival function differences between the 2 groups of tagged loggerheads (deep and light hooked) is not a strong inference because of the small sample size and the potential confounding of survey sampling design involving the use of different and long-duration duty cycles (see Fig. 3A).

Also, it is not clear whether it is possible to infer loggerhead post-release survival from the transmitter failure times used to derive the survival functions. The bathtub- or U-shape hazard functions (Fig. 2C) suggest that transmitter failure was a complex function comprising several competing risks or cause-specific failures (see discussion of this important issue in Lawless 1982 or Chiang 1991). The high early failures might have resulted from transmitter loss due to various component defects, improper attachment, or from early post-release mortality (see also Goodyear 2002, Graves et al. 2002 for similar comments regarding billfish tag-and-release studies). The early hazard or time-specific mortality rate is certainly much higher for the deep hooked group during this period (Fig. 2C), but loggerhead mortality and transmitter failure cannot be disentangled, so a reliable direct estimate of mortality for either group shortly after release is not possible.

Nonetheless, the KMT estimate of the deep hooked group failure rate during the first week after release was ca. 0.34 (95% complementary log-log confidence interval: 0.22 to 0.45) while the light hooked group estimate was ca. 0.08 (0.0 to 0.21). These KMT hazard or time-specific failure rates confound transmitter failure rates and loggerhead mortality if in fact any mortality did occur. Hence these estimates reflect at best the upper bounds on the apparent level of loggerhead mortality for the 2 groups shortly after release. These estimates must be viewed with extreme caution as they clearly reflect over-estimated failure and hence mortality probabilities attributable to capture in the longline gear. If all the transmitter failures during this period were due solely to equipment failures or tag loss then these estimates would not reflect loggerhead mortality in any way whatsoever.

The subsequent period of relatively constant time-specific mortality or failure (Fig. 2C) could reflect fewer failures, either because the early defective transmitters were already lost or that the likelihood of a turtle dying during this period had decreased: it is not possible to distinguish between these possibilities. The later accelerating period of time-specific failure or mortality could reflect the increasing likelihood of breakdown of the aging transmitters, the increasing likelihood of battery failure or the increasing likelihood of natural mortality of the turtle.

If the expected group-specific hazard functions reflect loggerhead time-specific mortality then it might be that deep hooked loggerheads were far more likely to die during the first 50 to 60 d after release from the longline gear than light hooked loggerheads. If the deep hooked loggerheads survived this long, there was apparently little difference thereafter between the survival chances of deep or light hooked loggerheads in this study, at least until around 200 d post-release. If this was the case then future studies using satellite telemetry might consider a 60 to 90 d sampling period as sufficient for estimating short-term survival or time-specific mortality for the released turtles. Battery power would not be an issue, and there would be no need to use long duty cycles such as 24/216, which serve only to considerably complicate any post-release survival study.

Using longer-term studies (>90 d) is not helpful, as the longer the time period, the greater the likelihood of confounding mortality risks, including the increasing risks of natural mortality and equipment failure (see also Goodyear 2002). Of course, transmitter failure during a short-term study does not ensure that it is possible to distinguish between early equipment failure from defects or mortality of the turtle from being hooked. The only way to distinguish between fishery-related loggerhead mortality and equipment failures, or loss due to improper attachment, is to determine the reason for each failure and to conduct a necropsy of all turtles that died while carrying a transmitter (see Work & Balazs 2002). Unfortunately, determining cause-specific failure at sea is not feasible (the turtles are not retrievable), although auxiliary information derived from time-depth recorder transmitters or perhaps pop-up tags could be useful here. This use of pop-up satellite tags was proposed by Graves et al. (2002) to study the post-release survival of billfish caught in recreational fisheries.

Another important issue that should be considered in future post-release survival studies concerns satellite tag assignment to a particular turtle retrieved alive from the longline gear. The NMFS observers were trained in many study protocols including how to identify the species of sea turtles retrieved from the longline gear, how to determine transmitter assignment and how to attach the satellite transmitter. We assume that there was no bias in the selection of a particular loggerhead for transmitter attachment because such bias could have a profound effect on estimation of the group-specific survival function. For instance, it is assumed that an observer did not assign transmitters to loggerheads that were thought likely to survive rather than assigning one to a turtle that was in poor condition when retrieved from the longline gear. This issue must be vigorously reinforced in any observer training program, just as it was in the present study.

Other major issues for consideration in future studies include issues such as what constitutes sufficient sample size and an appropriate experimental control for evaluation of post-release survival. While this was the largest data set of this type, comprising 40 satellite transmitters and a substantial observer program, it was nonetheless far too small to draw any strong inferences (see also Goodyear 2002 for similar conclusions regarding billfish tag-and-release studies). It is not known what a suitable sample size would be, but we estimate in the hundreds, even for an experiment comprising few treatment effects and no spatial or individual heterogeneity in turtle response to capture and release. A suitable and logistically feasible control for this type of quasi-experiment remains unclear.

Another important issue to consider in future studies is the potentially confounding problem of delayed or staggered entry into the study, since not all turtles were captured and released at the same time within each of the 4 sampling years. For instance, some turtles were caught and released in March, others in May, others in August, and so on. None of the transmitter failure times were right censored, so this is not an issue here. Delayed entry designs can be addressed using left truncation to account for variable entry time into the study (see Meeker & Escobar 1998 for detailed discussion of this issue), but should be avoided if possible. This is because such a study would involve an extremely complex and arbitrary censoring strategy possibly comprising left, right and interval censoring, as well as left truncation. Estimation of such a model would not be simple, and complex censoring design strategies seem pointless when it appears that the cause-specific hazards cannot be disentangled.

Recently, Hays et al. (2003) reported that various fisheries around the world account for an annual mortality probability of 0.31 for all sea turtle species exposed to those fisheries. Hays et al. (2003) based this estimate on satellite telemetry data. This study is problematic for many of the reasons discussed above for deriving satellite telemetry based mortality estimates. The Hays et al. (2003) study comprised no control, used tracking data that extended for more than 1 yr and so confounds natural mortality with any potential fishery-induced mortality, used transmitter failure as the basis for assigning a death to a particular turtle, failed to determine cause-specific failure for all transmitter failures and assumed all failures were due to turtle death, confounded duty cycle, transmitter types and research group methodologies, failed to use the transmission duration data to derive hazard functions, and used an extremely small sample size comprising 50 turtles from 3 species across all ocean basins with only 3 known deaths and so on. Hence the global annual mortality probability estimate of 0.31 of all sea turtle species exposed to fisheries that was proposed by Hays et al. (2003) is highly questionable at best.

Reliable estimates of natural and anthropogenic sources of mortality are known to be important for modelling the population or metapopulation dynamics of sea turtle stocks exposed to various hazards (Chaloupka 2002, 2003, 2004). While reliable estimates of natural mortality are becoming increasingly available for sea turtle populations (Chaloupka & Limpus 2002, Bjorndal et al. 2003, Seminoff et al. 2003), this is not the case for estimates of fishery-induced mortality. Satellite telemetry may be one useful approach for redressing this deficiency, but there are many challenges facing the use of this technology for robust evaluation of post-release survival of loggerhead sea turtles.

Acknowledgements. We are grateful to the NMFS observers who deployed all the transmitters on board the longline vessels. We thank Shawn Murakawa and Shandell Eames for technical assistance and observer training. We are grateful to Naoki Kamezaki and Colin Limpus for providing the data used in Fig. 1. We thank Bud Antonelis, Alan Bolten, Paul Dalzell, Peter Dutton, Francine Fiust, Judy Kendig, Mike Laurs, Jeff Polovina, Sam Pooley, Yonat Swimmer and Thierry Work for helpful comments on the manuscript. This work was supported by a NOAA National Marine Fisheries Service Contract to M.C.

LITERATURE CITED

Aguilar R, Mas J, Pastor X (1995) Impact of Spanish swordfish longline fisheries on the loggerhead sea turtle *Caretta caretta* population in the western Mediterranean. In: Richardson JI, Richardson TH (eds) Proceedings of the 12th annual symposium on sea turtle biology and conservation. NOAA Tech Mem NMFS-SEFSC-361, p 1–6

Balazs GH, Pooley SG (eds) (1994) Research plan to assess marine turtle hooking mortality: results of an expert workshop held in Honolulu, Hawaii, November 16–18, 1993. NOAA Tech Mem NMFS-SWFSC-201, p 1–166

Balazs GH, Pooley SG, Murakawa SKK (1995) Guidelines for handling marine turtles hooked or entangled in the Hawaii longline fishery: results of an expert workshop held in Honolulu, Hawaii, March 15–17, 1995. NOAA Tech Mem NMFS-SWFSC-222, p 1–41

Balazs GH, Miya RK, Beavers SC (1996) Procedures to attach a satellite transmitter to the carapace of an adult green turtle, *Chelonia mydas*. In: Keinath JA, Barnard DE, Musick JA, Bell BA (eds) Proceedings of the 15th annual symposium on sea turtle biology and conservation. NOAA Tech Mem NMFS-SEFSC-443, p 21–26

Bjorndal KA, Bolten AB, Reiwald B (1999) Development and use of satellite telemetry to estimate post-hooking mortality of marine turtles in the pelagic longline fisheries. Southwest Fish Sci Cent Admin Rep H-99-03C, Southwest Fisheries Science Center, National Marine Fisheries Service, Honolulu Lab, HI, p 1–25

Bjorndal KA, Bolten AB, Chaloupka MY (2003) Survival probability estimates for immature green turtles *Chelonia mydas* in the Bahamas. Mar Ecol Prog Ser 252:273–281

Bowen BW, Kamezaki N, Limpus CJ, Hughes GR, Meylan AB, Avise JC (1994) Global phylogeography of the loggerhead turtle (*Caretta caretta*) as indicated by mitochondrial DNA haplotypes. Evolution 48:1820–1828

Chaloupka M (2002) Stochastic simulation modelling of southern Great Barrier Reef green turtle population dynamics. Ecol Model 148:79–109

Chaloupka M (2003) Stochastic simulation modelling of loggerhead sea turtle population dynamics given exposure to competing mortality risks in the western south Pacific region. In: Bolten AB, Witherington BE (eds) Loggerhead sea turtles. Smithsonian Books, Washington, DC, p 274–294

Chaloupka M (2004) Exploring the metapopulation dynamics of the southern Great Barrier Reef green sea turtle stock and the possible consequences of sex-biased local harvesting. In: Akçakaya H, Burgman M, Kindvall O, Wood C, Sjogren-Gulve P, Hattfield J, McCarthy M (eds) Species conservation and management: case studies. Oxford University Press, New York, p 340–354

Chaloupka M, Limpus CJ (2001) Trends in the abundance of sea turtles resident in southern Great Barrier Reef waters. Biol Conserv 102:235–249

Chaloupka M, Limpus C (2002) Survival probability estimates for the endangered loggerhead sea turtle resident in southern Great Barrier Reef waters. Mar Biol 140:267–277

Chaloupka MY, Musick JA (1997) Age, growth and population dynamics. In: Lutz PJ, Musick JA (eds) The biology of sea turtles. CRC Press, Boca Raton, FL, p 233–276

Cheng IJ, Chen TH (1997) The incidental capture of 5 species of sea turtles by coastal setnet fisheries in the eastern waters of Taiwan. Biol Conserv 82:235–239

Chiang CL (1991) Competing risks in mortality analysis. Annu Rev Public Health 12:281–307

Cleveland WS (1993) Visualizing data. Hobart Press, Summit, NJ

Dutton PH, Balazs GH, Dizon AE (1997) Genetic stock identification of sea turtles caught in the Hawaii-based pelagic longline fishery. In: Epperly SP, Braun J (eds) Proceedings of the 17th annual symposium on sea turtle biology and conservation. NOAA Tech Mem NMFS-SEFSC-443, p 45–46

Gardner SC, Nichols WJ (2001) Assessment of sea turtle mortality rates in the Bahia Magdalena region, Baja California Sur, Mexico. Chelonian Conserv Biol 4:197–199

Glaser RE (1980) Bathtub and related failure rate characteristics. J Am Statist Assoc 75:667–672

Goodyear C (2002) Factors affecting robust estimates of the catch-and-release mortality using pop-off tag technology. Am Fish Soc Symp 30:172–179

Graves JE, Luckhurst BE, Prince ED (2002) An evaluation of pop-up satellite tags for estimating postrelease survival of blue marlin (*Makaira nigricans*) from a recreational fishery. Fish Bull 100:134–142

Hall MA, Alverson DL, Metuzals KI (2000) By-catch: problems and solutions. Mar Pollut Bull 41:204–219

Harrington DP, Fleming TR (1982) A class of rank test procedures for censored survival data. Biometrika 69:553–566

Hatase H, Takai N, Matsuzawa Y, Sakamoto W, Omuta K, Goto K, Arai N, Fujiwara T (2002) Size-related differences in feeding habitat use of adult female loggerhead turtles *Caretta caretta* around Japan determined by stable isotope analyses and satellite telemetry. Mar Ecol Prog Ser 233:272–282

Hays GC, Broderick AC, Godley BJ, Luschi P, Nichols WJ (2003) Satellite telemetry suggests high levels of fishing-induced mortality in marine turtles. Mar Ecol Prog Ser 262:305–309

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comp Graphic Statist 5:299–314

Julian F, Beeson M (1998) Estimates of marine mammal, turtle and seabird mortality for 2 California gillnet fisheries: 1990–1995. Fish Bull 96:271–284

Kamezaki N, Matsuzawa Y, Abe O, Asakawa H and 25 others (2003) Loggerhead turtle nesting in Japan. In: Bolten AB, Witherington BE (eds) Loggerhead sea turtles. Smithsonian Books, Washington, DC, p 210–217

Laurent L, Casale P, Bradal MN, Godley BJ and 14 others (1998) Molecular resolution of marine turtle stock composition in fishery bycatch: a case study in the Mediterranean. Mol Ecol 7:1529–1542

Lawless JF (1982) Statistical models and methods for lifetime data. John Wiley & Sons, New York

Limpus CJ, Limpus DJ (2003) The biology of the loggerhead turtle in western south Pacific Ocean foraging areas. In: Bolten AB, Witherington BE (eds) Loggerhead sea turtles. Smithsonian Books, Washington, DC, p 93–113

Loader C (1999) Local regression and likelihood. Springer-Verlag, New York

MathSoft (1999) SPLUS 2000 Guide to statistics, Vol 2. Data Analysis Products Division, MathSoft, Seattle, WA

McCracken ML (2000) Estimation of sea turtle take and mortality in the Hawaiian longline fisheries. Southwest Fish Sci Cent Admin Rep H-00-06. Southwest Fisheries Science Center, National Marine Fisheries Service, Honolulu Lab, HI, p 1–29

Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. John Wiley & Sons, New York

Paranjpe SP, Rajarshi MB (1986) Modelling non-monotonic survivorship data with bathtub distributions. Ecology 67: 1693–1695

Parker DM, Balazs GH, Murakawa SKK, Polovina JJ (in press) Post-hooking survival of sea turtles taken by pelagic longline fishing in the North Pacific. In: Coyne M (ed) Proceedings of the 21st annual symposium on sea turtle biology and conservation. NOAA Tech Mem NMFS-SEFSC

Poiner IR, Harris ANM (1996) The incidental capture, direct mortality and delayed mortality of turtles in Australia's northern prawn fishery. Mar Biol 125:813–825

Polovina JJ, Howell E, Parker DM, Balazs GH (2003a) Dive-depth distribution of loggerhead (*Caretta caretta*) and olive ridley (*Lepidochelys olivacea*) sea turtles in the central North Pacific: might deep longline sets catch fewer turtles? Fish Bull 101:189–193

Polovina JJ, Balazs GH, Howell EA, Parker DM, Seki MP, Dutton PH (2003b) Forage and migration habitat of loggerhead (*Caretta caretta*) and olive ridley (*Lepidochelys olivacea*) sea turtles in the central North Pacific Ocean. Fish Oceanogr 13:36–51

Sakamoto W, Bando T, Arai N, Baba N (1997) Migration paths of the adult female and male loggerhead turtles *Caretta caretta* determined through satellite telemetry. Fish Sci 63:547–552

Seminoff J, Jones T, Resendiz A, Nichols W, Chaloupka M (2003) Monitoring green turtles (*Chelonia mydas*) at a coastal foraging area in Baja California, Mexico: using multiple indices to describe population status. J Mar Biol Assoc UK 83:1355–1362

Slater J, Limpus C, Robins J, Pantus F, Chaloupka M (1998) Risk assessment of sea turtle capture in the Queensland east coast otter trawl fishery. Report prepared for TRAWL-MAC, Queensland Fish Management Authority on behalf of the Great Barrier Reef Marine Park Authority (Townsville, Queensland) and the Queensland Departments of Environment and Primary Industries (Brisbane, Queensland)

Turnbull BW (1976) The empirical distribution function with arbitrarily grouped, censored and truncated data. J R Statist Soc B 37:290–295

Work TM, Balazs GH (2002) Necropsy findings in sea turtles taken as bycatch in the North Pacific longline fishery. Fish Bull 100:876–880

Editorial responsibility: Otto Kinne (Editor),
Oldendorf/Luhe, Germany

Submitted: March 15, 2004; Accepted: July 13, 2004
Proofs received from author(s): September 21, 2004