DRAFT

Detection and distribution of significant clusters of Sea Lice (*Lepeophthericus salmonis* and *Caligus* sp.) infestation from samples of juvenile salmon and stickleback in the Broughton Archipelago, Knight Inlet, B.C. 2003-2006 using a spatial scan statistic (SaTScan™).

Introduction

Epidemiological researchers and Public Health professionals are often required to identify clusters, or areas of unusually high or low risk of infection, for further study or possible intervention, or to understand the underlying factors so that public health efforts can be focused on high risk areas. Similarly the outbreak of caligid ectoparasites *Lepeophthericus salmonis* and *Caligus* sp. commonly known as sea lice can be viewed as an infection and hence rates and occurrence of the outbreaks are subject to sampling and analysis. One method (the spatial scan statistic) appropriate to the identification of possible clusters of infestation is applied to infection rates of the salmon-specific caligid ectoparasite *Lepeophthericus salmonis* on out migrating juvenile Chum Salmon (*Onchorhynchus keta*) and Pink Salmon (*Oncorhynchus gorbuscha)* and resident Stickleback (*Gasterosteus aculaetus*) extensively sampled in the Broughton Archipelago and Knight Inlet of British Columbia by DFO (Department of Fisheries and Oceans Canada) from 2003 (Hargreaves *et al*, 2004) through 2006.

The Broughton Archipelago in British Columbia is a center for salmon aquaculture with 30+ sea farm sites located predominantly in the western portion of the chain (Figure 1). The Broughton and Knight Inlet also support large wild stocks of Chum (*Onchorhynchus keta*) and Pink Salmon (*Oncorhynchus gorbuscha)* which are important for both First Nations and commercial fisheries. Sea lice frequently infest farm salmon and numerous studies have linked sea lice infecting wild salmonids with the presence of sea farms (Tully and Whelan 1993, Costelloe *et al*. 1996, Todd *et al*, 1997, Mackenzie *et al*. 1998, Tully *et al* 1999, Bjorn *et al*, 2001, Bjorn and Finstad 2002, Marshall 2003,

Mortan and Williams 2004, Mortan *et al*. 2004, Mckibben and Hay 2004, Penston *et al* 2004, Carr and Whoriskey 2004, Krkosek *et al*. 2004). Strong concerns had previously been raised (e.g. Morton 2003; PFRCC 2003) about sea lice potentially originating from the numerous commercial salmon aquaculture farms located in the Broughton directly and negatively affecting sympatric wild populations of Chum (*Onchorhynchus keta*) and Pink Salmon (*Oncorhynchus gorbuscha)*. Indeed, commercial sea farms and subsequent lice infestation have been implicated in the collapse of Pink Salmon (*Oncorhynchus gorbuscha)* populations in the Broughton (Morton and Williams 2004, Morton *et al*. 2004).

In response to these concerns, on February 20, 2003 the Minister of Fisheries and Oceans Canada announced DFO's Pink Salmon Action Plan that focused on the Broughton Archipelago that included two major components: A marine monitoring program (MMP) whose objective was to obtain samples of juvenile pink salmon (*Oncorhynchus gorbuscha)* from the marine areas throughout the Broughton during the early sea life period to allow determination of the incidence and severity of infection by sea lice by location and time, and secondly, to regularly monitor the abundance of juvenile pink salmon at many locations during the early sea life period, to obtain additional information about the migration routes of juvenile pink salmon in the Broughton. While focused on juvenile pink salmon and their infection by parasitic sea lice (Calagid copepods), many other species of fish were also captured and sampled. Juvenile chum salmon (*Oncorhynchus keta)* were frequently found together in the same locations throughout the Broughton and were captured along with juvenile pink salmon (*Oncorhynchus gorbuscha)*. Juvenile pink and chum salmon also showed similar patterns

of infection by sea lice (Jones *et al*., 2004). Extensive data and analyses for both pink and chum salmon are presented in Hargreaves *et al* (2004) and Jones *et al* (2004). Due to the large volume of samples collected in the 2003-2006 MMP very few samples of other fish species have been fully analyzed in the laboratory, however, results were included for stickleback (*Gasterosteus aculaetus*) as high infection rates for this localized species was noted in the 2003 samples (Hargreaves *et al*, 2004).

The focus of this paper is to examine the distribution of infection rates of sea lice over time and space on migrating juvenile Chum (*Onchorhynchus keta*) and Pink Salmon (*Oncorhynchus gorbuscha),* and resident Stickleback (*Gasterosteus aculaetus*) using a spatial scan statistic (SaTScan^{TM}) (Kulldorff 2003, 2005).

Three primary questions are asked:

1) Are there statistically significant clusters of unusually high infection rates of sea lice in the sampling data for juvenile chum and pink salmon, and sticklebacks?

2) If significant clusters are found, how are they distributed in time and space?

3) If significant clusters are found, is there an association between the location of the significant clusters and locations of operational sea farms?

Methods

Data Sampling

Data used in this analysis was collected from the 2003-2006 MMP sampling program. The sampling methodology in the MMP is extensively described in Hargreaves *et al*. (2004) and is repeated here.

The overall scope of the DFO marine monitoring program in was designed to ensure data on juvenile pink salmon infection by sea lice were obtained over a broad area of the Broughton. Geographically, 106 sampling locations were selected. These sites included locations where pinks were captured as far away as practical from the salmon farms, locations where the pinks were in close proximity to salmon farms, and also, locations where pinks subsequently migrate through after passing by one or more of the salmon farms. To achieve these objectives, fish sampling was conducted at numerous locations within the Broughton (where most of the salmon farms were located) and to obtain samples as far away as practical from the salmon net pen sites, fish sampling was also conducted in Knight Inlet (Figure 1).

Knight Inlet is approximately 68 nautical miles long and the Klinaklini River enters the ocean at the head of Knight Inlet. This means that juvenile pink and chum salmon that originate in the Klinaklini River enter the ocean about 47 nautical miles distant from the nearest salmon farm (located at Sargeaunt Pass). This was the farthest distance that juvenile pink and chum salmon could be sampled that subsequently must eventually migrate past one or more of the salmon sea farms located in the Broughton. Another major stock of pink salmon originates from Glendale Creek, which enters the ocean about halfway along Knight Inlet (29 miles from the head of Knight Inlet), or about 17.5 nautical miles from the nearest salmon farm site at Sargeaunt Pass.

The specific sampling locations where chosen based mainly on visual inspection of the area by boat during the first week of March, 2003, combined with knowledge of the behaviour and migration patterns of juvenile pink and chum salmon. Six of the sampling locations that were chosen were also based on advice obtained from two local residents (Alexandra Morton and Billy Proctor, pers. comm.) regarding good locations for sampling juvenile pink and chum.

Each sampling location chosen for the MMP was carefully identified to ensure consistency in the sampling that was repeated every week. The latitude and longitude of each location was initially determined using Global Positioning System (GPS) marine navigation equipment. These data were written on the field data sheets used to record the fish catches and also entered into the vessel navigation course plotters using standard commercial computer software (Nobeltec Visual Navigation Suite). This position information was subsequently used each week to confirm the location for each sample obtained using the purse seine fishing gear in open water, or a beach seine on the shore. In addition, each location where samples were collected on the shore using beach seine fishing gear was also physically marked with both red spray paint (on the rocks or trees), and red flagging tape. While every attempt was made to collect fish samples at exactly these same locations every week, the precise position of some sets varied a small amount (typically less than 50 meters) from week to week. This was necessary because the variations in tide height, water currents, weather conditions, and logs, kelp debris and exposed rocks on some beaches made these locations temporarily very difficult to sample.

Some sampling locations were also subsequently eliminated, changed or added. A few sampling locations initially chosen during the first two weeks in March were subsequently eliminated. After the second week in March all of the sample locations in Zone A (near the head of Knight Inlet) were eliminated. This decision was made because too much sampling time was wasted by the purse seine vessel travelling from Port McNeill to the head of Knight Inlet each week. Although the Boston Whaler that was used for beach seining could travel much faster, the heavy sea conditions frequently encountered in Knight Inlet made it impractical and often dangerous to send the Boston Whaler too far ahead of the seine vessel. After the second week in March a few other sampling locations in both the Broughton and Knight Inlet were moved to new locations as close as possible (usually within 0.5 km) to the previous locations when it was found that the original locations could not be sampled under most conditions (e.g. at all tide heights). Some additional sampling locations were also added during April – June as the fishing operations and fish sampling procedures gradually became more efficient, which allowed more time for additional sets to be made each week.

To facilitate data analyses and comparison of results between different locations, each sampling location was numbered and the study area was divided into 11 "Zones" designated by the letters A through K (Figures 2 and 3). The location of the boundaries between these Zones was somewhat arbitrary but consistent throughout the study. For Knight Inlet the zone boundaries were chosen to divide the entire length of the Inlet into five sub-areas of roughly equal length. For the Broughton the criteria for choosing the Zone boundaries was based more on the location of natural physical divisions between major reaches and channels of water.

Samples of juvenile salmon and other fish were collected using two kinds of fishing gear: purse seines and beach seines. Both types of fishing gear have commonly been used for more than 50 years to sample juvenile salmon in scientific surveys. Seines were chosen because in the authors previous experience both these fishing gears allow more gentle capture and handling of juvenile salmon compared to some other gear types (e.g. trawl gear, troll gear, and gillnets) (e.g. Hargreaves *et al.* 1983, 1988). Seines also minimize the loss of scales from juvenile salmon and were also expected to eliminate or greatly reduce the loss of any attached sea lice.

Purse Seining

Purse seining is a standard fishing method employed typically in near-shore and coastal areas that is commonly used to capture both juvenile and adult salmon, and many other pelagic fish species (e.g. herring) in deeper water but it can also be a very effective method for capturing juvenile salmon in coastal areas when these fish are located in deeper water (e.g. depth > 5 m) but are also still relatively close to shore. Further, purse seining has also been used successfully to capture both juvenile and adult salmon in the open ocean (e.g. Hartt 1966; Pearcy and Fisher 1990).

For purse seining the fishing net is stored on a large hydraulically-rotated "drum" (spool) near the stern of the fishing vessel. To capture fish in a "circle set" the fishing vessel slowly moves forward in a circle pattern while the net is simultaneously unwound off the drum. The vessel circles completely around and when it reaches the end of the net that first entered the water, this end of the net is secured onto the side of the vessel. At this stage the net has enclosed a large body of water within the circle of the net, but the bottom of the net is still open. The open bottom of the net is then completely closed off

by using the hydraulic "purse winch" to gradually tighten a "purse line", which is a free running rope that passes through metal rings that are attached at frequent intervals near the bottom edge of the net. After the bottom of the net is closed off the net is ideally shaped like a cup or "purse" and hence the name "purse" seining. The net is then gradually spooled back onto the net drum until only a small portion of the net (known as the "bunt") still remains in the water alongside the vessel. This procedure gradually forces any fish caught inside the net to be crowded into a smaller and smaller amount of water. Finally all of the fish are forced into the bunt of the net alongside the vessel, where they can be easily brought aboard the vessel and retained (e.g. in commercial fisheries) or a sample taken and the rest of the fish released (e.g. in research programs). All of the purse seine sets completed in the MMP were "circle sets". No shore line was used and the nets were always "closed up" immediately after setting the net (i.e. the net was not held open to allow time for more fish to swim into the net before it was closed).

Three different purse seine nets were used for the DFO MMP in 2003. During the period from 2 March until 11 April the purse seine net used for sampling in the Broughton was 585 feet long and 54 feet deep. The stretched mesh size of the web in the lead section of this net was 3.2 cm (1.26 inches) and in the bunt section was 1.9 cm (0.75 inches). A liner panel of smaller mesh was attached inside the bunt section and had a stretched mesh size of 0.63 cm (0.25 inch). This lead section of webbing this net was older and the sewn seams in the bunt section of the net were weak and required frequent repairs. The corks on this net also provided bouyancy that was marginal, particularly in strong water currents. Therefore this net was replaced with another purse seine net on 13 April 2003. This replacement net was 183 meters (600 feet) long and 13 meters (42) feet

deep. The stretched mesh size was 2.54 cms (1 inch) in the lead section and 1.3 cms (0.5 inches) in the bunt section. This net also had a bunt liner panel identical to the previous net, with a stretched mesh size in the bunt of 0.63 mm (0.25 inches) that was attached to the inside of the bunt. These two purse seine nets used in the Broughton were very similar in length and depth, and the mesh of the bunt liners was identical. Therefore it is our opinion that it is very unlikely that the small differences between these two purse seine would have significantly affected the catches of juvenile pink or chum salmon.

A third net was used for all purse seine sampling that was conducted in Knight Inlet. This net was 274 meters (900 feet) long and 16.5 meters (54 feet) deep. The stretched mesh size of the web in the lead was 3.2 cm (1.25 inches) and in the bunt was 1.3 cm (0.5 inches). A liner panel of smaller mesh was also attached to the inside of the bunt of this net, and had a stretched mesh size of 0.63 cm (0.25 inches).

*****(Was similar equipment used for 2004-2006 ?– Check with Brent).

Several different vessels were used for purse seining in the MMP. All of these vessel masters had extensive (15-35 years) prior experience as skippers of purse seine vessels, although none had actually previously fished for specifically to capture juvenile salmon.

Beach Seining

Beach seining was the second type of fishing method used to obtain samples of juvenile salmon and other fish species in the 2003 MMP. Beach seining is a sampling method that is commonly used to capture juvenile salmon during the early sea life period, soon after the young salmon enter the marine environment. During this period juvenile pink and chum salmon are typically concentrated in the shallow water very close to shore

and capturing them by beach seining can be very effective. In both the Broughton and Knight Inlet, sampling with a beach seine was done by three or four people wearing chest waders, and using small open boats. Two Boston Whaler "Montauk" boats (17.5 and 18 foot lengths), each powered by a single 70 or 90 horsepower outboard motor, were used for most of the beach seining. An 18.5 foot long rigid-hull inflatable (Hurricane) was also used during March and as a substitute vessel when one of the Boston Whalers broke down.

To sample fish with the beach seine the net was loosely piled by hand into the bow of the small boats. At each chosen location the operator of the boat steered the boat slowly, bow first, into the shore until a second person could safely step ashore or into shallow water. This second person pulled this first end of the net ashore and then held fast this end of the net. The boat operator then slowly backed up the boat away from the shore, first perpendicular from the shore and then parallel to the shore line. As the boat moved along the beach the rest of the net come gradually came off the bow of the boat, usually without any assistance required from the people in the boat. As the last part of the net went into the water, the boat operator sharply turned the bow of the boat towards the shore and then slowly steered the boat directly into the shore until the third person could safely step ashore or into shallow water. This third person then pulled ashore the end of a long rope that was attached to second end of the beach seine net. The boat was then free of the net and was steered away from the net in the water. The third person then commenced to pull on the long rope until the second end of the net came ashore. Both people on the shore then slowly moved closer together along the shore while also continuing to steadily pull on their end of the net. Once the two ends of the net were

close together (about 6-10 feet apart) most of the net was gradually pulled in by hand and piled on the shore or in very shallow water. During this operation both the cork line and lead line were pulled simultaneously on each side of the net. This eventually crowded any fish caught in the net into the centre "bunt" portion of the net which had the smaller size mesh. As the last roughly 20- 30 feet of the net was pulled in to the shore the lead line was pulled slightly faster than the cork line. This resulted finally in the last section of the lead line coming out of the water while some of the net and cork line still remained in the water. When done correctly this caused the net to form a bag in shallow water in which all of the fish were retained and enclosed by the net, but the fish still remain fully submersed in water. All of the fish captured could then be easily removed from the net and then retained, sampled or released. After the catch was removed the net was manually piled back onto the bow of the boat and was ready for the next fishing operation.

It should be noted that the term "beach seine" may be a bit misleading. Sand or small stone beaches are rare throughout most of the Broughton and Knight Inlet area. With sufficient experience, skill and determination, however, it is possible to capture fish using beach seines at locations that commonly would not even be recognized as a "beach", including rocky shores and in some cases very steep shores or even shear cliffs. Many of the locations that were regularly sampled using beach seines in the MMP in 2003 in both the Broughton and Knight Inlet were of these more rugged types.
The nets used for beach seining in both the Broughton and in Knight Inlet were identical. Each net was 46 meters (150 feet) long and 3.7 meters (12 feet) deep, and was constructed of three 15.2 meter (50 foot) long panels sewn together. The web in the 15.2

m long panels at each end of the net had a stretched mesh size of 1.27 cm (0.25 inch). The middle 15.2 m long "bunt" panel of each net had a stretched web size of 0.64 cm (0.25 inch). A lead-filled rope ("lead-line"" was attached to the bottom edge of the net. Corks (99 total) were attached at regular intervals to a rope ("cork line") attached to the top edge of the net and provided sufficient buoyancy to float the entire net under all water current and weather conditions.

The purse seine vessels served as support vessels for the beach seining in both the Broughton and Knight Inlet. Each purse seine vessel had a skipper and a minimum of three additional crew members. The crew of the purse seine also conducted the beach seining. The routine for a typical day was for the skipper and all the crew to complete a purse seine set. Once the catch from this set was processed, a minimum of three of the crew members would immediately depart the purse seine vessel in the seine skiff (Boston Whaler) to go do a beach seine set at a location on the shore as nearby as possible to that purse seine location. Each beach seine set took about 30 minutes to complete and in the interim the skipper steered the purse seine vessel to the next purse seine set location. After the three man crew completed the beach seine set and processed the catch they caught up to the purse seine vessel with the faster Boston Whaler, and then all carried on to complete the next purse seine set. This pattern of alternating purse seine and beach seine sets continued all day. The skipper and all of the crew all slept and ate aboard the purse seine vessels. The purse seine vessels were totally self-sufficient for an entire week in the field. These vessels were re-provisioned in Port McNeill every Sunday afternoon, departed from Port McNeill every Monday morning, sampled all week in either the Broughton or Knight Inlet all week, and then returned to Port McNeill every Friday

evening. To maximize time available for sampling and minimize running time the purse seine vessels typically anchored in local protected waters each night close to wherever the last set beach or purse seine set of each day was made.

The purse seine vessels also served as support vessels for the beach seining in both the Broughton and Knight Inlet. Each purse seine vessel had a skipper and a minimum of three additional crew members. The crew of the purse seine also conducted the beach seining. The routine for a typical day was for the skipper and all the crew to complete a purse seine set. Once the catch from this set was processed, a minimum of three of the crew members would immediately depart the purse seine vessel in the seine skiff (Boston Whaler) to go do a beach seine set at a location on the shore as nearby as possible to the purse seine location. In the interim the skipper steered the purse seine vessel to the next purse seine set location. After the three man crew completed the beach seine set and processed the catch they caught up to the purse seine vessel with the faster Boston Whaler, and then all carried on to complete the next purse seine set. The skipper and all of the crew all slept and ate aboard the purse seine vessels. The purse seine vessels were re-provisioned in Port McNeil every Sunday and were otherwise self-sufficient for an entire week in the field. To maximize time available for sampling and minimize running time every day the purse seine vessels typically anchored in local protect waters each night close to wherever the last set beach or purse seine set of each day was made.
****(Need to make a statement here to indicate that the same procedures apply to the 2004-2006 sample seasons- Check with Brent)

While the above procedures were applied to all the sampling years, the start dates, end dates, and time frames were different for each year (Table 1).

****(Do we need to explain why this was so? – check with Brent)

Processing of fish samples

During the first two weeks of the MMP (2 – 14 March) in 2003 captured with either a beach seine or purse seine gear were examined for sea lice in the field. For sets in which small numbers of fish were caught (e.g. less than about 300), all fish were transferred using small fine mesh dip-nets directly from the bunt of the net into one or more five-gallon white plastic buckets filled with fresh seawater. For beach seining these buckets of fish were then loaded into the small boat used for beach seining and taken back to the larger purse seine vessel for processing. The transit time required to transport the fish in the buckets back to the purse seine vessels was typically five to ten minutes. Small numbers of fish caught with the purse seines were handled in a similar way: when the bunt of the purse seine net was alongside the vessel and mostly "dried up" all fish were gently removed using a long-handled dip-net and transferred directly into one or more five-gallon white plastic buckets. Once the fish were in these buckets they were anaethetised in the buckets by adding 1 to 5 mls of trimetasulfate (MS222), depending on the volume of seawater in each bucket. Each fish was then individually examined by manually picking it up and visually inspecting both sides of the fish to identify the fish species and count any sea lice that were visible on the fish. Some fish were also examined for sea lice under a dissecting microscope. The fork length (for salmonids) or total length (for non-salmonids) of each fish was then measured. During this process a random sample (e.g. every 10th fish that was examined) of 30 fish of each species was

retained by placing each fish in an individual plastic Whirl-Pak© or Zip-Lock© bag, depending on the size of the fish. After removing the sample of 30 any remaining fish were placed in additional plastic five-gallon buckets filled with fresh seawater until they recovered from the anaesthesia and then released back into the ocean. The samples of fish that were retained in plastic bags were labelled and immediately placed in a nine cubic foot 120 volt A.C. chest freezer.

The procedures described above that were used to process the fish captured in the MMP were substantially modified beginning in mid-March. By the end of the second week of field sampling it became apparent that large catches of fish could not be fully processed in the field without severely slowing down the sampling program. In addition, during the first week of March the lead author invited Alexandra Morton to observe the sampling protocol that was being used in the MMP. She expressed concerns that the procedures that were being used to sorting, transport, handle and sample juvenile salmon in the field might result in some motile sea lice being lost. As a result, this protocol of measuring fish and examining them for sea lice in the field would be terminated. A new procedure was implemented, for both beach seine and purse seine catches, in which fish were placed into individual Whirl-Pak© or Zip-Lock© bags directly from the bunt of the nets. Once the net had been dried up sufficiently to concentrate the fish in the bunt, samples of up to 30 fish of each species were removed directly from the bunt, one fish at a time. Each fish in the bunt was captured alive into an individual sample bag. Care was taken to avoid or minimize handling each fish prior to capturing it into the individual sample bag. The fish that were individually bagged were also chosen as randomly as possible from the entire catch in the net. However, it should be recognized that this

sampling procedure was not truly random (e.g. removing all fish caught one fish at a time and retaining every 10th fish), especially when catches were large. Capturing each fish into an individual bag plastic bag typically resulted in a small amount of seawater also being enclosed by the bag. To minimize the time required to freeze the fish each bag was pricked several times with a sharp needle to allow the seawater to drain out of the bag. The diameter of the needles was much smaller than the size of even the smallest motile sea lice, to ensure no sea lice that might fall off the fish would be lost through the holes in the bags. The non-motile stages of sea lice are smaller (down to microscopic size) but are firmly attached to the fish with a strong filament. It was assumed that these younger and smaller stages of sea lice would remain attached to the fish and therefore not be lost through the holes in the bags.

The samples of fish that were retained in plastic bags were labelled and immediately placed in a nine cubic foot 120 volt A.C. chest freezer aboard the purse seine vessels. These samples remained in the freezers aboard the purse seine vessels until the end of each week. On Friday each week the fish samples were transferred from the freezers on the purse seine vessels into large "Coleman" coolers equipped with pre-frozen "freezer packs". These coolers were then taken by vehicle from Port McNeill to the DFO Pacific Biological Station (PBS) in Nanaimo, B.C., where the fish samples were transferred from the coolers into large plastic garbage bags and then stored in a walk-in freezer at –20 degrees Celsius.

The frozen fish samples that were stored in the freezer at the PBS were subsequently individually examined in the fish health laboratory at PBS to confirm the species identifications and the length measurements recorded in the field. Each fish was also

examined visually and with a dissecting or compound microscope to count the number of sea lice on each fish. Samples of sea lice were subsequently also examined microscopically to identify the species and development stages of the sea lice. The analysis of the fish samples and sea lice was initiated in March 2003 at the Pacific Biological Station. Two or three people worked continuously full-time on this task every week until the end of December 2003. By that time all of the juvenile pink and chum salmon samples that were collected in the MMP in 2003 had been analyzed. However, due to the very large number of juvenile pink and chum samples that were collected only a small portion of the samples of other fish species, mainly three-spine stickleback and Pacific herring) had also been analyzed by early March 2004.

When the total catches of fish from each set were small (e.g. less than 300 fish per species), the samples of 30 fish per species were immediately bagged and then all of the remaining fish from each set were identified by species, counted and released. When the catches were larger (e.g. 300 - 500) typically one or two people bagged fish, while the other person(s) simultaneously counted, identified and released fish from the bunt. For very large catches (e.g. >500), fish were also removed from the bunt using dip-nets, rather than individually. In these cases the total number of dip-nets of fish that were removed was counted, and all the fish in several dip-nets were randomly chosen (e.g. every fourth dip-net of fish if there were 12 dip-nets of fish in total) and all the fish removed in the dip-net were placed in five gallon white buckets. After all of the fish in the bunt had been removed, the fish in these buckets were individually counted and identified. The total numbers of fish of each species that were originally captured in the seine net was then estimated by multiplying the total number of dip-nets of fish that were

removed from the bunt by the average number of fish of each species in the dip-net samples that had been retained in the white buckets. In some cases the catches of herring were very large and it was impractical to count every fish or even by dip-netting the fish out of the bunt. In these cases the catch of herring by weight (tonnes) was estimated visually by the purse seine vessel skipper, and this weight was subsequently converted to number of fish by dividing the weight in tonnes by an estimated average weight per fish.

These protocols were used in subsequent sampling years.

Data Management

In the field the catch from each set of a beach seine or purse seine net was recorded manually on a separate page of paper referred to as the "field data sheets" (FDS). These FDS consisted of 8.5 by 11 inch water-proof paper that was pre-printed with a data template that included blanks for the set location, time, date, and catches for various common fish species. These FDS were collated aboard the purse seine vessels each day and returned each week to the Pacific Biological Station (PBS) in Nanaimo, B.C. each week along with the frozen fish samples.

At the PBS the data recorded on the field data sheets were manually entered into a computer spreadsheet (MS Excel). Each week the updated spreadsheet was transmitted by DFO internal email to a database manager at the DFO Institute of Ocean Sciences in Sidney, B.C. The database manager added the new data each week to a relational database and conducted various manual and automated checks to verify the accuracy and quality of the data. For example, the data for the location of each beach seine and purse seine was checked to ensure the GPS coordinates that were recorded in the field were

actually within the boundaries of the study area and closely matched the actual coordinates and name of the standard sampling locations.

After all the field work was completed and the data from both the field data sheets and the analyses of fish samples at PBS had been loaded into the relational database at IOS (Institute of Ocean Science), the data were completely verified again by comparing the data originally entered on the FDS and laboratory data sheets with the electronic data in the relational database at IOS. Subsets of the data were transferred to an Access® database for subsequent data extraction and analysis.

Data Adjustments and Corrections

The data verification procedures described above identified a variety of data errors. In many cases these could be traced to human errors made in recording or transcribing data. For example, the locations of some fishing sets were obviously incorrect (e.g. entirely out of the sampling region, or up on land rather than in the water) and by comparison with the coordinates for that standard sampling location these could be traced back to simple errors made in recording the GPS coordinates. In other cases the time or date of the set was obviously incorrect, when compared with bridge logs of the purse seine vessels or the time sequence of other sets that were completed that same day. These types of obvious errors were simply corrected in the final relational database, and these corrections were also noted on the original field data sheets.

The original catch data from the field were also corrected and adjusted based on the laboratory results. Juvenile chum salmon are easily distinguished from juvenile pink salmon soon after these species enter the marine environment (e.g. by the highly visible

"parr" marks (dark coloured vertical bands) on chum, and subsequently by the smaller size of scales in pinks). As the fish grow older and larger, however, pink and chum can become increasingly difficult to distinguish by using only the external characteristics. Eventually additional internal features, such as the number and appearance of the gill rakers, must frequently be examined to confirm the species identifications. Capturing each fish alive into an individual plastic bag meant that these internal characteristics could not be examined without removing the fish from the bag and handling them. In this study this was not desirable because of the risk of losing sea lice from either the fish or the sample bag. Therefore the species identifications in the field were done as best they could with the fish remaining inside the plastic bag. It was recognized that this would result in some errors in species identification, and therefore also some errors in the original field catch data. To correct for these errors the final catch data in the relational database were adjusted based on the species identifications that were later confirmed in the analyses of the frozen fish samples in the laboratory at the PBS. For example, assume that the original field catch data indicated that 200 chum and 200 pink were caught in a particular beach or purse seine set, and that 30 fish of each species were bagged and frozen. If the subsequent analyses of these frozen samples at PBS indicated that 15 of these "pinks" were actually chum, then the original catch data were adjusted proportionally to 300 chum and 100 pinks captured.

Spatial Scan Statistic

The spatial scan statistic is a method to support the detection and inference for the spatial clustering of disease (Kulldorff and Nagarwalla 1995, Kulldorff 1997). The test is designed to detect variably sized clusters located anywhere in the study region (Kulldorff 1997). Kulldorff implemented the test statistic in SaTScan™ software (Kulldorff 2005) which can be used to analyze spatial, temporal, and space-time point sampling data. The software is designed to: 1) evaluate reported spatial or space-time disease clusters and to determine if they are statistically significant, 2) evaluate geographic surveillance of disease to detect areas of significantly low or high rates, 3) to test whether a disease is randomly distributed over space or over time or over space and time and 4) to facilitate early detection of disease outbreaks through repeated time-periodic disease surveillance.

Although the spatial scan statistic was designed to investigate clusters in epidemiological investigations, the implementation of the spatial scan statistic analysis in SaTScan™ has been used in a wide variety of fields. For example: (1) Infectious Diseases (Cousens *et al.* 2001, Fevre *et al.* 2001, Chaput and Heimer 2002, Huillard d'Aignaux *et al.* 2002, Mostashari *et al.* 2003, Sauders *et al.* 2003), (2) Cancer Research (Hjalmars *et al.* 1996, 1999, Kulldorff *et al.* 1997, Imai 1998, VanEenwyk *et al.* 1999, Viel *et al.* 2000, Sheehan *et al.* 2001, Gregorio *et al.* 2001, Roche *et al.* 2002, Jemal *et al.* 2002, Michelozzi *et al.* 2002, Thomas and Carlin 2003, Gregorio and Samociuk 2003, Buntinx *et al.* 2003), (3) Pediatrics (Sankoh *et al.* 2001, George *et al.* 2001, Imai 1998, Forand *et al.* 2002), (4) Sclerosis (Walsh and Fenster 1997, Sabel *et al.* 2003), (5) Lupus (Walsh and DeChello 2001), (6) Diabetes (Green *et al.* 2003), (7) Alcohol and Drug

dependencies (Hanson and Wieczorek 2002), (8) Veterinary Medicine (Norström *et al.* 2000, Ward 2001, USDA 2001, Doherr *et al.* 2001, Perez *et al.* 2002, Schwermer *et al.* 2002, Enemark *et al.* 2002, Ward 2002, Falconi *et al.* 2002, Knuesel *et al.* 2003), (9) Wildlife Veterinary Medicine (Smith *et al.* 2000, Berke *et al.* 2002, Miller *et al.* 2002, Hoar *et al.* 2003, Olea-Popelka *et al.* 2003), (10) Forestry (Coulston and Riitters 2003), (11) Toxicology (Sudakin *et al.* 2002), (12) Psychology (Margai and Henry 2003), (13) Brain Imaging (Yoshida and Miyashita 2003), and (14) Criminology (Jefferis 1998, Kaminski *et al.* 2000). We use the spatial scan statistic to evaluate our first two primary questions as stated above.

Data input into SaTScan™ is either dichotomous disease information such as in case-control data or counts such as the number of cases among a population at risk in a geographic area. The user can select either a binomial distribution (as would be appropriate for case-control data) or a Poisson distribution (as would be appropriate for count data). The program adjusts for the underlying heterogeneity of a background population and for count data using the Poisson model, SaTScan™ can adjust for any number of categorical covariates as comparisons are often made between area and therefore rates or risk estimates need to be adjusted for confounding variables. The current version (5.1.3) of SaTScan™ also allows a space-time permutation model (Kulldorff 2003, 2004) that requires only case data, and information about the spatial location and time for each case. The number of observed cases in a cluster is compared to what would have been expected if the spatial and temporal locations of all cases were independent of each other so that there is no space-time interaction.

SaTScan™ software and documentation are available for downloading free of charge from the National Cancer Institute's (NCI) web site (Kulldoff 2005). Currently, SaTScan™ is sponsored by the Statistical Research and Applications Branch with the NCI Division for Cancer Control and Population Science's Surveillance Research Program.

Spatial scan statistics such as SaTScan™ have the unique ability to specifically detect clusters and test their significance. Neither the cluster size nor the regions need to be specifically defined in advance. The evaluation of clearly defined null and alternative hypothesis follows from the derived test statistic which is based on a likelihood ratio and not on an ad hoc procedure. Coultson and Riitters (2003) further point out that "the test is valid regardless of the actual spatial pattern and the approach works with data at multiple spatial scales".

Details of the computation of the test statistic for the Bernoulli and Poisson models are outlined in Kulldorff (1997) while the computational details of the Permutation model is described in Kulldorff *et al.* (2004). Briefly, for the Binomial and Poisson models, there is a requirement that either a uniform population at risk is met through out the study region, or that other "denominator" data that provides information about the population at risk is available. For example, in epimediological research, underlining census population numbers are useful denominator data that provide accurate expected number of cases based on the underlying population (Kulldorff *et al.* 2004). Let *i* represent an index of a set of units of interest that define a study area, then Si represents

a geographic location for index *i*, usually, the center of a geographic area. Let M_i represent the size of a population of interest in unit *i*, and let N_i represent the number of individuals in the population M_i that have some attribute of interest. An event is the occurance of the attribute of interest.The objective of the scan statistic is to identify clusters of the measured units (S*i*) for which the occurrence of the attribute of interest is significantly more likely within the cluster than outside of the cluster.

The procedure to determine significant clusters begins with examining each location (S*i*) in the study area. At each (S*i*) circular windows of different sizes are imposed with the S*i* at the center of the windows. It is possible that a window may contain different S*i*'s such that there may be n_1 measurement units and n_2 windows imposed upon each unit. The total number of windows in the study area then equals n_1 * n_2. Each window potentially contains different sets of neighbouring units, and each is a potential cluster. A likelihood ratio is then used to determine the significance of a potential cluster. The number of events in each measurement unit (*i*) is assumed to be Bernoulli and Poisson distributed (Kulldoff 1997). The test statistic then for a specific window *w* is defined by the likelihood ratio under the null hypothesis that rates of events within the window is the same as rates of the events everywhere else (Kulldorff 1997, Kulldorff *et al.*1997) and is calculated as:

$$L_w := \left[\frac{\left[\left(\frac{N_{ci}}{M_{ci}} \right)^{N_{ci}} \cdot \left(\frac{N_{co}}{M_{co}} \right)^{N_{co}} \right]}{\left[\left(\frac{N^t}{M^t} \right)^{N^t} \right]} \right] \qquad (1)$$

where N and M refer to the number of events and population size respectively. The subscripts ci and co refer to the totals of those variables over the measurement units within (ci) and outside (co) of the window. $M^T = \sum_i M_i$ is the total population size in the study area and $N^T = \sum_i N_i$ is the total number of events in the study area. To set up a one way test of the null hypothesis against the alternative that the rate of events is higher within the window, the statistic L_w is multiplied by an indicator function *I*. *I* has a value of 1 if $N_{ci}/M_{ci} > N_{co}/M_{co}$ and zero otherwise. The likelihood for a specific window w is proportional to:

$$\left[\left(\frac{N_{ci}}{\mu}\right)^{N_{ci}} \cdot \left[\frac{\left(N^t - N_{ci}\right)}{\left(N^t - \mu\right)}\right]^{\left(N^t - N_{ci}\right)}\right] \cdot I \qquad (2)$$

where μ is the expected number of events witih the window under the null hypothesis that the rate of events is the same across the study area. *I* is an indicator function that in this case has a value of 1 if $N_{ci} > \mu$ and zero otherwise. The ratio's N_{ci}/μ and $(N^t - N_{ci}) / (N^t - \mu)$ are proportional to the ratios within and outside the window respectfully, and for fixed N^t and μ, the likelihood increases with increased number of events in the window N_{ci}.

For all windows w, the L_w's are ranked and the corresponding window with the highest maximum likelihood ratio (maxium L_w) is the primary or most likely cluster. The other clusters are termed secondary clusters. The distribution of the primary L_w cluster and simulated P value are determined using Monte Carlo simulation that replicates the analysis for a large number of random replications of the original data set under the null hypothesis of complete randomness of clusters (Kulldorff 1997). The significance of the

primary cluster L_w is determined by comparing it's value to the distribution of L for all windows from the Monte Carlo simulation and if L_w is greater than 95% of the values form the simulation the cluster is considered significant at the 5% level. The significance of secondary clusters is determined in the same way however, Kulldorff (1997) points out that the values of these estimates are to be considered approximate and conservative.

The procedure is extended to three dimensions by the use of cylinders for scanning rather than windows. The height of the cylinder represents time while the base of the cylinder represents the space. The scanning proceeds by allowing the base and height to vary continuously as the scan progresses through space and time. The calculation of the likelihood ratio or the significance test is not changed.

For the space-time permutation model a slightly different approach is needed (Kulldorff *et al.*2004). A requirement of nearly all scan statistics is the necessity of either a uniform population at risk, a control group, or some other denominator data the provides information about the population at risk (Kulldorff *et al.*2004). The scanning under this model utilizes a very large collection of cylinders to define the scanning window with each window being a possible candidate for an outbreak or cluster. As with the other models, the base of the cylinder represents the special location while the height of the cylinder represents time. The time scale is set by the user and in the case of the sea lice data, it is weekly (7 days). The computation of the primary cluster is based solely on the calculations based on case data. The computation is done as follows (Kulldorff *et al.*2004). Suppose that there are event or case counts C for s locations during time t such that there are a set of C_{st}. Then the total number of observed cases is:

$$C = \sum_s \sum_t C_{st} \tag{3}$$

hence, the expected number of cases μ_{st} can be calculated as:

$$\mu_{st} = \left(\frac{1}{C}\right) * \left(\sum_s C_{st}\right) * \left(\sum_t C_{st}\right) \tag{4}$$

which is the proportion of all cases that occurred in area s times the total number of cases during time t. The expected number of cases for a particular cylinder A is the summation of all μ_{st} over all the time periods within that cylinder:

$$\mu_A = \left(\sum_{(s,t)\in A} \mu_{st}\right) \tag{5}$$

Kulldorff *et al.*(2004) show that if we let C_A be the observed number of cases in the cylinder then when μ_A is small compared to the total number of observed cases C, then C_A is approximately Poisson distributed with mean μ_A when there is no space-time interactions. This approximation leads to a Poisson likelihood as a measure that cylinder A contains a cluster. This is calculated as:

$$L_A = \left(\frac{C_A}{\mu_A}\right)^{C_A} \left(\frac{C - C_A}{C - \mu_A}\right)^{(C - C_A)} \tag{6}$$

Kulldorff *et al.*(2004) state that "this is the observed divided by the expected to the power of the observed *inside the cylinder*, multiplied by the observed divided bye the expected to the power of the observed *outside the cylinder*. The cylinder with the maximum likelihood constitutes the space-time cluster of cases that is least likely to be a chance occurrence, and hence, the primary candidate for a true outbreak."

Evaluation of statistical significance is done using Monte Carlo hypothesis testing as stated earlier, however, as there is no denominator data to adjust the case data, the dates and time of the cases are randomly recombined while ensuring that the marginals of the spatial and temporal components are unchanged. (See Kulldorff *et al.*(2004) for details).

Recently, the space-time permutation model has been used for the syndromic surveillance for outbreaks of West Nile Virus (WNV) (Mostashari *et al.*2003) where mosquitoes collected from set traps in New York city where examined for NWV. As the traps were fixed as to location, it was assumed that trapping effort was consistent among the traps. In this type of study, it is important that trapping effort is consistent (Kulldorff pers. comm.). We believe that the set sampling locations and consistent methodology used for sea lice data collected in the Broughton Archipelago ensured that the effort was similar among locations. The catch per unit effort and catch effort for the Broughton and Knight Inlet data show remarkable consistency lending support to the view that the data from each sampling location is consistent with minimal location bias (Hargreaves *et al* 2004).

Data for each species was extracted from the database for each year and converted to dbf file formats for input into SaTScan^{TM} . For each species and year, a typical file set includes: a location or Co-ordinate file that contains the identifier, name, and latitude, longitude for each of the sample sites used in that year, and a Case File that contains the location identifier for each sample and the number of fish examined in the sample and the number of fish infected with sea lice in that sample. As the analysis was done using the

Permutation model, a Control file (for the Bernoulli model) and a Population file (for the Poisson model) were not needed.

Distance Analysis

As part of the overall examination of sea lice infection rates, we plotted for each sample week within a sample year, the abundance of lice (lice/fish) for each sea lice life stage for Chum and Pink salmon and for stickleback, for each sample location landward for specific sea farms. These specific farms were active fish farms that were geographically determined to be the first farm that migrating salmon fry would encounter as they moved from their natal streams toward the sea. The first farm (site 25) was located in Sargeaunt's Pass, Tribune Channel and the second farm (site 5) was located in the Sutlej Channel of Kingcome inlet. For each farm location, the distance was set as 0 and the distance in kilometres to each sample location within 30 kilometres landward and seaward for site 5 and 90 kilometres landward and 30 kilometres seaward for farm site 25 was measured using direct minimum line over water from the specific fish farm to the closest sample site first and then to subsequent sample sites. The abundance of lice, measured as lice/fish for each life stage was plotted against these normalized distances from the two specific sea farm sites. The pattern of lice abundance and proximity to significant clusters from the spatial analysis over the normalized distance was then examined.

Where significant clusters occurred, we also examined where possible, the lice abundance records for the nearest active fish farm.

Results

Discussion

References

Berke O, von Keyserlingk M, Broll S, Kreienbrock L. 2002. On the distribution of Echinococcus multilocularis in red foxes in Lower Saxony: identification of a high risk area by spatial epidemiological cluster analysis. Berliner und Munchener Tierarztliche Wochenschrift. 115:428-434.

Buntinx F, Geys H, Lousbergh D, Broeders G, Cloes E, Dhollander D, Op De Beeck L, Vanden Brande J, Van Waes A, Molenberghs G. 2003. Geographical differences in cancer incidence in the Belgian province of Limburg. European Journal of Cancer, 39:2058-72.

Chaput EK, Meek JI, Heimer R. 2002. Spatial analysis of human granulocytic ehrlichiosis near Lyme, Connecticut. Emerging Infectious Diseases, 8:943-948.

Coulston JW, Riitters KH. 2003. Geographic Analysis of Forest Health Indicators Using Spatial Scan Statistics. Environmental Management, , 31: 764-773.

Cousens S, Smith PG, Ward H, et al. 2001 Geographical distribution of variant Creutzfeldt-Jakob disease in Great Britain, 1994-2000. The Lancet, 357:1002-1007.

Doherr MG, Hett AR, Rufenacht J, Zurbriggen A, Heim D. 2002. Geographical clustering of cases of bovine spongiform encephalopathy (BSE) born in Switzerland after the feed ban. Veterinary Record, 151: 467-472.

Enemark HL, Ahrens P, Juel CD, Petersen E, Petersen RF, Andersen JS, Lind P, Thamsborg SM. 2002. Molecular characterization of Danish Cryptosporidium parvum isolates. Parasitology, 125:331-341.

Falconi F, Ochs H, Deplazes P. 2002. Serological cross-sectional survey of psoroptic sheep scab in Switzerland. Veterinary Parasitology, 109:119-127.

Fevre EM, Coleman PG, Odit M, et al. 2001. The origins of a new Trypanosoma brucei rhodesiense sleeping sickness outbreak in eastern Uganda. The Lancet, 358:625-628.

Forand SP, Talbot TO, Druschel C, Cross PK. 2002. Data quality and the spatial analysis of disease rates: congenital malformations in New York State. Health and Place, 8:191-199.

George M, Wiklund L, Aastrup M, Pousette J, Thunholm B, Saldeen T, Wernroth L, Zaren B, Holmberg L. 2001. Incidence and geographical distribution of sudden infant death syndrome in relation to content of nitrate in drinking water and groundwater levels. European Journal of Clinical Investigation, 31: 1083-1094.

Green C, Hoppa RD, Young TK, Blanchard JF. 2003. Geographic analysis of diabetes prevalence in an urban area, Social Science and Medicine, 57:551-560.

Gregorio DI, Kulldorff M, Barry L, Samociuk H, Zarfos K. 2001 Geographic differences in primary therapy for early stage breast cancer. Annals of Surgical Oncology, 2001; 8:844-849.

Gregorio DI, Samociuk H. 2003 Breast cancer surveillance using gridded population units, Connecticut, 1992 to 1995. Annals of Epidemiology, 13:42-49.

Hamilton, L.C. 1992. Regression with Graphics: A second course in Applied Statistics. Belmont California: Duxbury Press. ISBN: 0-534-15900-1

Hartt, A. C. 1966. Migrations of salmon in the north Pacific Ocean and Bering Sea as determined by seining and tagging. 1950-1960. Bull. Int. North Pac. Fish. Comm. 19. 141p.

Hanson CE, Wieczorek WF. 2002. Alcohol mortality: a comparison of spatial clustering methods. Social Science and Medicine, 55:791-802.

Hargreaves, N. B., B. A. Patten and R. J. LeBrasseur. 1983. Beach seine catches of juvenile salmon and other fish in Masset Inlet, Masset Sound, and Juskatla Inlet, B.C. in 1983. Can. Dta Rep. Fish Aquat. Sci. 683. 267 p.

Hargreaves, N. B., B. A. Patten, E. W. Carter, and R. J. Le Brasseur. 1988. Two-boat trawl, gillnet, purse seine, angling, fish trap, and fry trap catches of salmonids and other fishes captured in Masset Nlet, B.C. in 1983, 1984, and 1985. Can. Data Rep. Fish. Aquat. Sci. 687. 177 p.

Hargreaves, N. B., D. Herriott, and R. V. Palermo 2004. Pink Salmon Action Plan: Abundance and distribution of juvenile salmon and other fish caught in the BroughtonArchipelago, Knight Inlet and Muchalat Inlet, B.C. in 2003. In Press.

Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. 1996. Childhood leukemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. Statistics in Medicine, 15:707-715.

Hjalmars U, Kulldorff M, Wahlquist Y, Lannering B. 1999. Increased incidence rates but no space-time clustering of childhood malignant brain tumors in Sweden. Cancer, 85:2077-2090.

Hoar BR, Chomel BB, Rolfe DL, Chang CC, Fritz CL, Sacks BN, Carpenter TE. 2003. Spatial analysis of Yersinia pestis and Bartonella vinsonii subsp berkhoffii seroprevalence in California coyotes (Canis latrans). Preventive Veterinary Medicine, 56:299-311.

Huillard d'Aignaux J, Cousens SN, Delasnerie-Laupretre N, Brandel JP, Salomon D, Laplanche JL, Hauw JJ, Alperovitch A. 2002. Analysis of the geographical distribution of sporadic Creutzfeldt-Jakob disease in France between 1992 and 1998. International Journal of Epidemiology, 31: 490-495.

Imai J. Spatial disease clustering in Kochi prefecture in Japan. 1998 National Institute of Public Health Epidemiology and Biostatistics Research, 57-96 (in Japanese).

Jefferis ES. 1998. A multi-method exploration of crime hot spots: SaTScan results. National Institute of Justice, Crime Mapping Research Center.

Jemal A, Kulldorff M, Devesa SS, Hayes RB, Fraumeni JF. 2002. A geographic analysis of prostate cancer mortality in the United States. International Journal of Cancer, 101:168-174.

Judd, C. M. & McClelland, G. H. 1989. Data analysis: A model-comparison approach. New York: Harcourt Brace Jovanovich. ISBN: 0-15-516765-0

Kaminski RJ, Jefferis ES, Chanhatasilpa C. 2000. A spatial analysis of American police killed in the line of duty. In Turnbull et al. (eds.), Atlas of crime: Mapping the criminal landscape. Phoenix, AZ: Oryx Press.

Knuesel R, Segner H, Wahli T. 2003. A survey of viral diseases in farmed and feral salmonids in Switzerland. Journal of Fish Diseases, 26:167-182.

Kulldorff M, Nagarwalla N. 1995 Spatial disease clusters: Detection and Inference. Statistics in Medicine, 14:799-810.

Kulldorff M. A spatial scan statistic. 1997. Communications in Statistics: Theory and Methods, 26:1481-1496.

Kulldorff M, Feuer EJ, Miller BA, Freedman LS. 1997. Breast cancer in northeastern United States: A geographical analysis. American Journal of Epidemiology, 146:161-170.

Kulldorff M, Athas W, Feuer E, Miller B, Key C. 1998. Evaluating cluster alarms: A space-time scan statistic and brain cancer in Los Alamos. American Journal of Public Health, 88:1377-1380.

Kulldorff M. 1999. Spatial scan statistics: Models, calculations and applications. In alakrishnan and Glaz (eds), Recent Advances on Scan Statistics and Applications. Birkhauser.

Kulldorff M. 2001. Prospective time-periodic geographical disease surveillance using a scan statistic. Journal of the Royal Statistical Society, A164:61-72.

Kulldorff M, Heffernan R., Hartman J, Assunção R.M., Mostashari F. 2004. A space-time permutation scan statistic for the early detection of disease outbreaks. Manuscript, 2004.

Kulldorff M. and Information Management Services, Inc. SaTScan™ v6.0: Sofrware for the spatial and space-time scan statstics. http://www.SaTScan.org/ , 2005. SaTScan™ is a trademark of Martin Kulldorf. The SaTScan™ software was developed under the joint auspices of Martin Kulldoff, of the National Cancer Institute and of Farzad Mostashari at the New York Cuity Department of Health and Mental Hygiene.

Margai F, Henry N. 2003. A community-based assessment of learning disabilities using environmental and contextual risk factors. Social Science and Medicine, 56: 1073-1085.

Michelozzi P, Capon A, Kirchmayer U, Forastiere F, Biggeri A, Barca A, Perucci CA. 2002. Adult and childhood leukemia near a high-power radio station in Rome, Italy. American Journal of Epidemiology, 155:1096-1103.

Miller MA, Gardner IA, Kreuder C, Paradies DM, Worcester KR, Jessup DA, Dodd E, Harris MD, Ames JA, Packham AE, Conrad PA. 2002. Coastal freshwater runoff is a risk factor for Toxoplasma gondii infection of southern sea otters (Enhydra lutris nereis). International Journal for Parasitology, 32:997-1006

Mostashari F, Kulldorff M, Hartman JJ, Miller J, Kulasekera V. 2003. Dead bird clustering: A potential early warning system for West Nile virus activity. Emerging Infectious Diseases. Vol 9. No 6.

Norström M, Pfeiffer DU, Jarp J. 2000. A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds. Preventive Veterinary Medicine, 47: 107-119.

Olea-Popelka FJ, Griffin JM, Collins JD, McGrath G, Martin SW. 2003. Bovine tuberculosis in badgers in four areas in Ireland: does tuberculosis cluster? Preventive Veterinary Medicine, 59:103-111

Pearcy, W. G. and J. P. Fisher. 1990. Distribution and abundance of juvenile salmiondsoff Oregoan and Washington, 1981-1985. NOAA Tech. Rep. NMFS 93

Perez AM, Ward MP, Torres P, Ritacco V. 2002. Use of spatial statistics and monitoring data to identify clustering of bovine tuberculosis in Argentina. Preventive Veterinary Medicine, 56: 63-74.

Roche LM, Skinner R, Weinstein RB. 2002. Use of a geographic information system to identify and characterize areas with high proportions of distant stage breast cancer. Journal of Public Health Management and Practice, 8:26-32.

Sabel CE, Boyle PJ, Loytonen M, Gatrell AC, Jokelainen M, Flowerdew R, Maasilta P. 2003. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. American Journal of Epidemiology, 157: 898-905.

Sankoh OA, Ye Y, Sauerborn R, Muller O, Becher H. 2001. Clustering of childhood mortality in rural Burkina Faso. International Journal of Epidemiology, 30:485-492.

Sauders BD, Fortes ED, Morse DL, Dumas N, Kiehlbauch JA, Schukken Y, Hibbs JR, Wiedmann M. 2003. Molecular subtyping to detect human listeriosis clusters. Emerging Infectious Diseases, 9:672-680.

Schwermer H, Rufenacht J, Doherr MG, Heim D. 2002. Geographic distribution of BSE in Switzerland. Schweizer Archiv fur Tierheilkunde, 144:701-708.

Sheehan TJ, Gershman ST, MacDougal L, Danley R, Mrosszczyk M, Sorensen AM, Kulldorff M. 2001. Geographical surveillance of breast cancer screening by tracts, towns and zip codes. Journal of Public Health Management and Practice, 6: 48-57.

Smith KL, DeVos V, Bryden H, Price LB, Hugh-Jones ME, Keim P. 2000. Bacillus anthracis diversity in Kruger National Park. Journal of Clinical Microbiology, 38:3780-3784.

Sudakin DL, Horowitz Z, Giffin S. 2002. Regional variation in the incidence of symptomatic pesticide exposures: Applications of geographic information systems. Journal of Toxicology - Clinical Toxicology, 40:767-773.

Thomas AJ, Carlin BP. 2003 Late detection of breast and colorectal cancer in Minnesota counties: an application of spatial smoothing and clustering. Statistics in Medicine, 22:113-127.

United States Department of Agriculture. West Nile virus in equids in the Northeastern United States in 2000. USDA, APHIS, Veterinary Services, 2001.

VanEenwyk J, Bensley L, McBride D, Hoskins R, Solet D, McKeeman Brown A, Topiwala H, Richter A, Clark R. 1999. Addressing community health concerns around SeaTac Airport: Second Report. Washington State Department of Health, 1999.

Viel JF, Arveux P, Baverel J, Cahn JY. 2000. Soft-tissue sarcoma and non-Hodgkin's lymphoma clusters around a municipal solid waste incinerator with high dioxin emission levels. American Journal of Epidemiology, 152:13-19.

Walsh SJ, Fenster JR. 1997. Geographical clustering of mortality from systemic sclerosis in the Southeastern United States, 1981-90. Journal of Rheumatology, 24:2348-2352.

Walsh SJ, DeChello LM. 2001. Geographical variation in mortality from systemic lupus erythematosus in the United States. Lupus, 10:637-646.

Ward MP. 2001. Blowfly strike in sheep flocks as an example of the use of a time-space scan statistic to control confounding. Preventive Veterinary Medicine, 49: 61-69.

Ward MP. 2002. Clustering of reported cases of leptospirosis among dogs in the United States and Canada. Preventive Veterinary Medicine, 56:215-226.

Yoshida M, Naya Y, Miyashita Y. 2003. Anatomical organization of forward fiber projections from area TE to perirhinal neurons representing visual long-term memory in monkeys. Proceedings of the National Acedemy of Sciences of the United States of America, 100:4257-4262.