Using the H-index to assess disease priorities for salmon aquaculture

Alexander G Murray[a], Maya Wardeh[b,c] and K Marie McIntyre[b,c]

[a]Marine Laboratory Marine Scotland Science, 375 Victoria Road, Aberdeen AB11 9DB, UK

[b]Health Protection Research Unit in Emerging and Zoonotic Infections, University of Liverpool, L69 7BE, UK

[c] Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, Leahurst Campus, Chester High Road, Neston, Cheshire CH64 7TE, UK

Abstract

Atlantic salmon's (*Salmo salar*) annual aquaculture production exceeds 2M tonnes globally, and for the UK forms the largest single food export. However, aquaculture production is negatively affected by a range of different diseases and parasites. Effort to control pathogens should be focused on those which are most "important" to aquaculture. It is difficult to specify what makes a pathogen important; this is particularly true in the aquatic sector where data capture systems are less developed than for human or terrestrial animal diseases. Mortality levels might be one indicator, but these can cause a range of different problems such as persistent endemic losses, occasional large epidemics or control/treatment costs. Economic and multi-criteria decision methods can incorporate this range of impacts, however these have not been consistently applied to aquaculture and the quantity and quality of data required is large, so their potential for comparing aquatic pathogens is currently limited. A method that has been developed and applied to both human and terrestrial animal diseases is the analysis of published scientific literature using the H-index method. We applied this method to salmon pathogens using Web of Science searches for 23 pathogens. The top 3 H-indices were obtained for: sea lice, furunculosis, and infectious salmon anaemia; post 2000, Amoebic Gill Disease (AGD) replaced furunculosis. The number of publications per year describing bacterial disease declined significantly, while those for viruses and sea lice increased significantly. This reflects effective bacterial control by vaccination, while problems related to viruses and sea lice have increased. H-indices by country reflected different national concerns (e.g. AGD ranked top for Australia). Averaged national H-indices for salmon diseases tend to increase with log of salmon production; countries with H-Indices significantly below the trend line have suffered particularly large disease losses. The H-index method, supported by other literature analyses, is consistent with the nature and history of salmon diseases and so provides a useful quantitative measure for comparing different diseases in the absence of other measures.

Key words: Atlantic salmon, pathogens, aquaculture, H-index

Highlights:

Ranking pathogens of salmon aquaculture is difficult

We use publication trends and H-index to rank pathogens

Sea lice, infectious salmon anaemia, furunculosis have highest H-indices

Bacterial publications in decline, reflects vaccination

Virus and sea lice publications increasing reflecting emerging problems

1. Introduction

Aquaculture is an increasingly important source of protein and now accounts for approximately 50% of fish used for human consumption (FAO 2014). In cooler temperate marine waters the Atlantic salmon (*Salmo salar*) is the principle farmed species. Salmon farming began in the 1960s in Norway and soon after was established in other countries such as Chile, Canada and the Faeroe Islands. Global production now exceeds 2M tonnes (FAO 2014). However, production has been impacted by diseases and parasites, for example large outbreaks of infectious salmon anaemia in Chile (Mardones et al. 2011) and on-going high costs of sea lice control around the world (Costello 2009a).

To effectively target disease control, or implement risk reduction programmes, it is necessary to identify which diseases are “important”. A range of such methods are used for targeting diseases in humans and terrestrial animals (Brooks et al. 2015). Assessment is complicated because diseases have different types of impact upon production e.g. causing mortalities, reduced productivity, treatment costs, or loss of employment. In aquatic animals, assessment is particularly difficult given a lack of data and the need to assess a range of impacts in different circumstances and in different countries. Selection can be made using expert opinion (Murray 2015), but an objective selection based on impact would be more meaningful and defensible (Brooks et al. 2015).

The most obvious disease impact is mortality, but its economic consequences are different depending on whether death occurs early or late in the production cycle, as losses of full grown fish are more costly than those of young smolts (Kilburn et al. 2012). Mortalities can occur as occasional large epidemic shocks or near constant endemic losses; these can be budgeted for within normal production costs. Sea lice may cause only limited mortality in well-run fish farms (Soares et al. 2011) but their treatment imposes large costs on salmonid farmers (Costello 2009a) and lice may have impacts on third parties, as elevated burdens can be found on wild fish up to 30 km from farms (Middlemas et al. 2012). Other diseases, such as infectious salmon anaemia (ISA) only impact farmed fish, but costs under area control strategies applied to ISA can fall on neighbours (Murray et al. 2010). The uncertainties caused by epidemics include serious social costs such as short-term loss of jobs and an uncertain investment climate that prevents creation of new employment. Diseases can also cause loss of potential production by limiting scope for aquaculture e.g. in Australia, amoebic gill disease (AGD) limits marine salmon farming to areas with good access to freshwater used to treat the pathogen.

Impacts are therefore multifactorial and analysis tools such as multi-criteria-decision-analysis, MCDA (Del Rio Vilas et al. 2013, Brooks et al. 2015), or different impacts turned into an economic cost and compared between diseases, can be useful. These approaches allow diseases to be ranked, for example such as has been undertaken for exotic pig diseases in Australia (Brooks et al. 2014). Both approaches need good characterisation of different impacts to allow multiple diseases to be compared in a consistent manner; for aquatic animal diseases data may be absent and systematic assessment methods are consequently less consistently applied. Economic estimates for individual disease impacts are often made e.g. for ISA (Hastings et al. 1999, Mardones et al. 2011), sea lice (Costello 2009a), or piscirickettsia (Rozas and Enriquez 2014). However most costings are based on expert opinion or limited calculations and relatively few use systematic and transparent methods e.g. pancreas disease (Aunsmo et al. 2010) or bacterial kidney disease (Hall et al. 2014). Fofana and Baulcomb (2012) have applied an economic model to assess the costs of three different diseases and this approach may be an area of progress in the near future. However, even if systematic approaches are available, lack of data limits the ability to make detailed economic assessment in many cases. Some costs, such as to welfare, are very difficult to assess, and although methods such as contingency valuation do exist these have severe limitations in practice (Venkatachalam 2004). Opinions on significance of impacts vary depending on different stakeholder's concerns (Brooks et al. 2014). It is therefore very difficult, using existing methods, to compare the economic impact of different salmonid diseases in a consistent way.

The scientific literature is, by its very nature, well documented. Academic publications provide a measure of the effort, and therefore importance, that scientists and their funders attach to different diseases, and continuity in publications on a topic suggests an ongoing issue. An approach that has been utilised for comparing the significance of different diseases is to use scientific publications as a proxy for interest and hence effectively an assessment of the importance of pathogens.

Specifically, citation histories can be generated and linked to years of publication, subject areas, countries and so forth. A particularly good summary measure of citation is the Hirsch or H-index method for each particular disease (McIntyre et al. 2011, 2014a). In this case, the H-index is the integer at which the number of papers equals the number of citations arising as a result of those papers.

The H-index approach has been applied to diseases of humans and domestic animals (McIntyre et al. 2014a). It is an objective measure that has been shown to be related to a combination of morbidity and mortality effects (via Disability-Adjusted Life Years) that result from these diseases (McIntyre et al. 2011, 2014a, Cox et al.2016). Although certain diseases may attract disproportionate interest, or lack of interest (such as neglected diseases in lower income countries (Hunter 2009)), the literature is a good descriptor of the effort placed in preventing or controlling particular diseases. While the scientific literature and H-index method are not immune to biases, they can be used for objective comparison of interest in diseases. The use of the H-index method has considerable potential to facilitate the development of measures that allow policy and industry assessment of generalised disease control strategies to be focused on those diseases that have the highest H-indices, rather than focussed on a subjective selection of diseases of interest. Within this study, the H-index method was used to objectively assess and compare the interest in aquaculture diseases in salmon producing countries, as a measure to help identify disease priorities.

Countries are affected by different diseases to differing extents, for example Australian salmon farms are heavily affected by AGD but, at least until recently, this has been a lesser problem in most other countries. Conversely, Australia lacks many of the diseases that cause serious problems in other countries. For the major salmon pathogens, the potential effects of changes in the focus of scientific research with time were also examined.

2. Methods

2.1. H-index literature search protocol

2.1.1. Information sources

H-index searches were undertaken in May 2015 using Web of Science (WoS) (WOS, 2014) and the methods described by McIntyre et al. (2014a). Previous work has established that results of H-index searches for pathogens undertaken using different bibliographic sources (e.g. WoS, SCOPUS, Google Scholar) are not identical but are highly correlated (McIntyre et al. 2011).

2.1.2. Eligibility criteria

Searches were restricted to the years 1990 to 2014, inclusive. The effects of time were examined for nine salmon pathogens by calculating H-indices based on searches spanning both 1990-2014 and 2000-2014, inclusive. English is used in WoS, however searches also include foreign-language publication title translations. All literature in the WoS database has been published.

2.1.3. Searches

Searches were undertaken using search phrases specified in quotation marks (“”), and the ‘topic’ (TS – examining the full paper) or ‘title’ (TI)’ search field and with no lemmatization. Phrases were compiled including pathogen scientific name, alternative names, synonyms and alternative spellings according to NCBI Taxonomy (NCBI, 2014). H-index scores for clinical diseases used clinical terms as well as pathogen phrases for the main pathogens of disease. Virus searches also included synonyms and acronyms from the NCBI Taxonomy database (NCBI, 2014) and International Committee on Taxonomy of Viruses (ICTV, 2014), and the term ‘virus’, and excluded other entities (viral or non-viral) which shared acronyms. The Boolean operators ‘AND’, ‘OR’, and ‘NOT’ linked multiple search phrases. The full search terms were generated within the Enhanced Infectious Disease Database (EID2) (McIntyre et al. 2014b, Wardeh et al. 2015). All searches were carried out on the same day (6^{th} May 2015) to avoid biases in time to publication.

2.1.4. Search phrases

The literature searches took the format of:

TS = ((Host) AND (Pathogen OR Disease))

Or TI = ((Host) AND (Pathogen OR Disease))

Analysis was also broken down by country using queries of the format:

TI = ((Host) AND (Pathogen OR Disease)) AND CU = (Country)

where TI was a Title search and TS was a topic search (a general search of the full paper).

2.2. Hosts

The host considered in analyses was the Atlantic salmon (*Salmo salar*). Search phrases used both English and Linnaean names. The unqualified term “salmon” is sometimes used to refer to Atlantic salmon and so searches allowed “salmon”, while explicitly excluding other species of salmon. Also, “infectious salmon” was excluded to avoid papers on infectious salmon anaemia in other host species. The host search term was thus:

("Salmo salar" OR "Atlantic salmon" OR (salmon NOT (“coho salmon” OR “Chinook salmon” or “pink salmon” OR “chum salmon” OR “Sockeye salmon” OR “Masu salmon” OR “king salmon” OR “Pacific salmon” OR “infectious salmon”))).

2.3. Diseases and pathogens

Compiling the list of diseases and pathogens of importance to salmon farming is complicated as many are known by a variety of names. The Scottish government maintains a list of diseases of interest to Scottish aquaculture (http://www.scotland.gov.uk/Topics/marine/Fish-Shellfish/aquaculture/diseases), which was used as a basic source of diseases perceived to be of significance to salmon farming. This list consisted of 22 diseases of finfish, however one of these was spring viraemia of carp, which was not of relevance. Another listed condition was cataracts, which does not relate to a specific pathogen and so was excluded; *Diplostomum spathaceum*, a specific pathogen causing cataracts was, however, included. This left 20 diseases for analyses. To ensure more global coverage, a further three salmonid diseases were added: heart and skeletal muscle inflammation (HSMI) is a disease of increasing concern in Norway, Epizootic haematopoietic necrosis (EHN) is a disease of trout that is notifiable to the OIE, and salmon rickettsia syndrome (SRS) is an important disease in Chile that also occurs in the UK. The initial list of pathogens was extensive; those with the highest H-index scores were selected for more detailed analyses later. The incorporation of an initial long-list that included relatively obscure salmon pathogens ensured that a final list of key pathogens with high H-indices was not dependent on this initial list.

Multiple pathogens may be associated with disease conditions, and publications can use different names for the same pathogen owing to taxonomic revision, or to disagreement. Sometimes, no pathogen has yet been associated with a particular disease. Situations when multiple, or zero, pathogens are named in association with particular diseases are described in Table 1.

2.4 Analysis of Data

2.4.1. Comparison of title and text search methods

The TI and TS searches generated different H-index scores. These sets of scores were compared by calculating a regression coefficient and r^2 value, and computing confidence intervals. Specific diseases for which H-indices lay outside the confidence range were identified and used to characterise the outcomes of the TI and TS searches.

Nine key diseases with the highest H-index scores (>10) were selected for further analysis.

2.4.2. Variation in H-index scores between countries

Search phrases for the salmon pathogens with the highest H-indices were broken down by country to examine national differences in interest in the nine key diseases with the highest world-wide H-index scores. Eight countries were incorporated in this analysis including: Norway, Chile, Scotland, Canada, Denmark (The Faeroe Islands), Australia, USA and Ireland. These countries accounted for 99.5% of world salmon production in 2012 (FAO 2014). Mean national H-indices were calculated and plotted against national production data for 2012. A regression with 95% confidence intervals was used to identify countries in which averaged H-indices deviated significantly from the normal.

2.4.3 Changes in disease publications with time

Diseases emerged and declined with time. For those with the top nine H-index scores calculated using publications from 1990-2014, H-indices were recalculated using only publications from the 2000-2014 period, to identify pathogens whose scores changed.

A more detailed analysis of change in the pathogens' profiles with time was carried out by plotting the annual number of publications, as identified by the TI search, for each of the nine key diseases. A regression analysis was used to identify statistically significant trends with time. All statistical analyses were undertaken in R (Crawley 2013).

3. Results

3.1. Comparison of title and text search methods

The H-indices for salmon calculated using full text (TS) searches ranged from 61 for typical frunculosis to 0 for Red Mark Syndrome and for the title only (TI) searches they ranged from 39 for sea lice to 0 for several conditions (Table 2). H-index scores calculated for TI were much lower than for TS.

If the H-index scores calculated using TI and TS searches were plotted against each other (Fig. 1), there was a good degree of agreement between values (r^2 of 0.77), suggesting that both searches provide similar results in terms of the rankings of specific pathogens. The TI search was a more specific indicator of papers describing salmon-pathogen interactions, and therefore these were used for the remaining analyses.

The H-index method was used to identify key diseases or parasites for salmon: sea lice, furunculosis, ISA, IPN, PD, AGD, *G. salaris*, Vibrio and BKD (Table 2). All these diseases had, for Atlantic salmon, H-index scores in excess of 10 for searches by paper title (TI). Further analyses in this paper focuses on these diseases/parasites.

3.2. Variation in H-index scores between countries

The results suggest that the focus of research on diseases is dependent on country (Table 3): Scotland, Canada, Chile and the USA all had their highest H-indices for sea lice, while for Norway the highest was for furunculosis, with ISA a close second. For Denmark the highest H-indices was for *G. salaris*, for Australia it was AGD and for Ireland it was PD. Norway also had a substantial H-index score for *G. salaris* and both Norway and Scotland for IPN. Canada had its highest national H-indices for ISA and BKD; the latter possibly due to the widespread BKD reservoir in wild Pacific salmon spilling over into farmed Atlantic salmon.

The calculation of national H-index scores for diseases (Table 2) allowed examination of the presence of disease being associated with a non-zero H-index score at the country level. Australia is free of all the diseases listed in Table 2, except AGD and vibriosis (Munday et al.1992). PD has not been reported in Australia, Canada, Chile, Denmark or the USA (OIE 2015), while Gs is absent from Scotland, Ireland and non-European countries (OIE 2015). Ireland is additionally free of BKD and ISA (apart from a single event of ISAV detection in trout). Of 21 zero H-index scores (Table 2), seven are associated with the presence of disease (odds 1:2) and for the 51 non-zero H-indices, disease is present in 45 countries (odds 15:2). A Fisher's exact test gave an odds ratio of 0.070 with a 95% confidence range of 0.016 to 0.270 ($p = 7.1 \times 10^{-6}$), confirming that overall, a national focus of research is associated with the presence of disease.

If the mean was taken of national H-index scores for the main nine diseases/pathogens, the result was correlated with the logarithm of salmon industry production; the regression line had an r^2=0.340 or r^2=0.713 if Chile was excluded (Fig 2). There is thus a general relationship between salmon production and mean H-index score, with exceptions for Denmark and Chile.

3.3. Changes in disease publications with time

The incidences of diseases of significance are likely to change with time. When H-indices were re-calculated to exclude references published before 2000, then two of the bacterial diseases (BKD new H-index score = 5 and vibriosis new H-index score = 8) dropped out of the top nine diseases. The list of highest-ranking pathogens became: sea lice = 34, ISA = 26, AGD = 22, IPN = 22, furunculosis = 18, PD = 17, *G. salaris* = 12, HSMI =10 and CMS = 9. This revised list was dominated by viruses (5 of the top 9 H-indices), but the H-indices for sea lice and AGD were higher than those for most other conditions.

To assess changes in publication rates with time in more detail for the nine key diseases, the numbers of papers identified using TI searches by year were examined (Fig 3). The analysis showed statistically significant declines in publication rates for two of the three bacterial diseases (vibriosis and furunculosis, Fig. 3a) and significant increases in publication rates for all three viral diseases (IPN, ISA and PD, Fig 3b). There was a significant increase in publications concerning sea lice, while the other parasitic conditions (AGD and Gs) showed peaks in publications in the middle of the 2000-2010 decade, and did not have significant trends over the whole period (Fig. 3c).

4. Discussion

4.1. Optimising search methods: TS versus TI

For the analysis of the H-index score, we used H-indices calculated from the results of papers searched by their titles only, TI, as opposed to more general searches using the entire contents of the paper, TS, to identify publications containing pathogen and host search terms.

The TS search identified more papers in which salmon was mentioned in a general way, as opposed to being the host in specific host-pathogen interactions. The bacterial diseases furunculosis and vibriosis were above the 95% confidence range of the regression line of H-indices obtained using TI and TS searches and BKD lay on this upper confidence limit (Fig. 1). This may be because these pathogens affect a broad range of hosts and many papers identified under TS searches describe infections occurring in non-salmonids or in terms of their general mechanisms of infection. In these cases, terms relating to salmon are not likely to appear in the title but, since they are key diseases of salmon, they are likely to be discussed within the text. The TS search thus over-estimates the significance of these bacterial diseases as specific conditions of salmon. Similarly VHS, IHN and ERM which are serious diseases of trout, lie above the regression line, perhaps as the focus of most papers is on trout (mention of salmon may occur in the discussion or as keywords) (Fig 1).

Conversely, the diseases below the regression line's 95% confidence range were PD, AGD, HSMI and CMS, with IPN on the lower confidence bound (Fig. 1). These are likely to be closely linked with the species affected and so appear in the paper's title. CMS and HSMI are restricted to salmon and although AGD and IPN affect other species, they are only of major concern when infecting salmon. Therefore the TI searches identified a relatively high proportion of the papers that TS searches identified, and as these pathogens are of most concern to salmon, this is desirable.

ISA lies almost exactly on the regression line, however ISA publications could be underestimated in a title search, as authors may not explicitly specify salmon as the host for infectious salmon anaemia (Mardones et al. 2011, Murray et al. 2010). This may mean that the H-index score of ISA in salmon calculated using a TI search is an under-estimate of its significance. However the position on the regression line means TI and TS searches would give this pathogen similar rank, so use of TS would not correct any bias in publications.

In the citation index there are several papers associated with the key words of "Atlantic Salmon" but whose content makes no or negligible reference to these species. For example Metzger et al. (2010) is a paper on *R. salmoniarum* in Chinook salmon; this paper contains no mention of Atlantic Salmon, except within the titles of two papers listed in its bibliography, however "Atlantic Salmon" is listed as a "keyword plus" in Web of Science. These papers are detected using a TS search, but excluded by a TI search; the TI search is thus more appropriate.

In conclusion, the use of title searches (TI) is more specific than full text searches (TS), which may include papers on pathogens that are not relevant to salmon. The narrow TI search focuses on papers that concern the pathogen's interaction with salmon, rather than general papers on the pathogen, ensuring that publications are relevant. The TI search excludes some relevant papers, but as the aim is to use the H-index method as a tool for comparison rather than to calculate it for its own sake, this is not relevant except where the exclusion is biased.

4.2. Key Diseases identified from the H-index analysis

The highest H-index scores were used to identify a list of key diseases or parasites of salmon: sea lice, furunculosis, ISA, IPN, PD, AGD, *G. salaris*, Vibrio and BKD. All had TI H-indices of >10. This is a taxonomically diverse range of pathogens including viral diseases (ISA, IPN and PD), bacterial diseases (furunculosis, vibriosis and BKD) and parasites (AGD, *G. salaris* and sea lice).

4.2.1. Variation between countries and relationship to salmon production

H-indices for most countries relate to the national level of salmon production; they do not relate to a countries' population or gross national product (GNP) (regressions of H-index scores against these variables gave insignificant results). This indicates that the H-index method is identifying relevant research related to national salmon production activity rather than just to a nation's economic or scientific resources, and it can therefore justly be used as an indicator of the scale of different disease problems.

There are two notable exceptions, Chile and Denmark, for which H-indices lay below the lower 95% confidence limit of the regression (Fig. 2). Chile is a developing country with fewer resources to pay for research (Hunter 2009), however investment in Chilean aquaculture science has increased substantially in recent years, leading to increased activity. The annual number of Chilean publications on salmon (TI = Salmon AND cu = Chile) shows a significant increase with time and in recent years the numbers of publications have approached those of Scotland (Fig. 4). The salmon production of Denmark actually occurs in the self-governing Faeroes. The small population of these islands (50,000) limits resources for research and collaboration may consequently occur with Norway or Scotland as much as Denmark proper. Therefore the H-index scores for Denmark are smaller than expected given the level of production. It is perhaps worth noting that both the Faeroes and Chile have experienced major collapses in their salmon production with falls of, respectively, 40 thousand tonnes (kt) to 13 kt (2004-6) and 388 kt to 120 kt (2008-10) (FAO 2012); both drops of approximately 70% and both associated with ISA epidemics.

Non-zero H-indices are statistically significantly associated with the presence of disease in a country, suggesting that H-index scores are more related to a country's Atlantic salmon production than to, for example, its population or GNP, and further endorsing the use of this method to characterise national pathogen problems.

4.2.2. Changes in disease publications with time

Diseases emerge and decline as aquaculture and disease control methods develop, and so publication patterns and H-index scores would also be expected to change with time.

4.2.2.1 *Bacterial diseases*

There has been a major decline in numbers of publications describing the key bacterial diseases furunculosis and vibriosis, although not BKD (Fig. 4a). Historically, control of bacterial diseases has been a major concern since the early 20th Century (Mackie 1935). The decline follows the introduction of effective vaccines for these diseases (Håstein et al. 2005) and is also reflected in a major drop in antibiotic use (Alderman and Hasting 1998). Effective vaccines do not yet exist for BKD and this pathogen is controlled in salmon in Scotland using movement restrictions (Murray et al. 2012).

The rickettsial disease SRS is a relatively minor issue in most salmon farming countries and so does not rank by H-index as a key disease (H-index = 8), but it is an extremely serious problem in Chile, with direct losses estimated at over $100M (Rozas and Henriquez 2014). The H-index score for SRS publications from Chile is 3 (not shown), which would make it the second highest for that country. Chile cannot rely on research occurring in the northern countries to control this disease and has continued to use more antibiotics than other salmon producers. Increasing Chilean research (Fig. 4) into SRS would be expected to provide nationally-specific answers.

4.2.2.2 *Viral diseases*

Research output on the key viral diseases, ISA, IPN and PD, has increased over the last 25 years (Fig 3b) although with some oscillation, such as peaks of PD publications in 1995-8 and again in recent years. In Scotland, PD was a major issue in the 1990s, and re-emerged to a new peak of losses in 2006-7 (Kilburn et al. 2012). It is currently of considerable interest in Norway, owing to large-scale regional control policies (Tavornpanich et al. 2012). Infectious salmon anaemia has had peaks of publication in the late 1990s with its emergence in Norway, in 2000 following a major outbreak in Scotland, after 2007 with subsequent worldwide spread including recurrence in Scotland (Murray et al. 2010), and a very large outbreak in Chile (Mardones et al 2011). All these events have contributed to research interest and hence publications that increase with disease episodes. Publications on IPN have also increased over the 1990-2014 period as the pathogen has become widespread (Murray 2006) and caused substantial mortality in salmon smolts (Kilburn et al. 2012), although vaccines are now reducing impacts.

4.2.2.3 *Parasites*

The rise of published sea lice research has been spectacular (Fig 3c), with numbers of publications comparable to the sum of those for all viral diseases. However year-on-year changes have been extremely variable; this is in spite of the large numbers that would be expected to dampen stochastic variation with a normal distribution. Inter-annual variation reflects the nature of the sea lice research community, which hold dedicated international sea lice conferences at which much of their work is presented. This differs substantially from viral or bacterial research, which is presented at a range of general virology (or bacteriology), fish pathology, aquaculture or epidemiology conferences. As a result of dedicated conferences, the sea lice topic has focussed years of publication, with 2000 being exceptional because the conference journals for both a 3rd and 4th international conference were published in that same year. Other peaks, such as 2002, 2009 and 2011 also reflect publication of conference special issues of journals. The spikes in publication therefore reflect the sociology of scientists rather than the epidemiology of lice. However, the trend mirrors a clearly increasing lice problem associated with increasing aquaculture density and reducing efficacy of medicines. Sea lice are currently the most serious limitation to expansion of salmon production almost wherever salmon are farmed (Jones and Beamish 2011), with the exception of Australia. Sea lice are also the subject of a growing controversy concerning their impact on wild salmonids (Costello 2009b).

Publications on Amoebic Gill Disease and *Gyrodactylus salaris* both peaked around 2005 and research has since declined (Fig 3c). However AGD publication rate increased significantly over the period of 1990-2014, partly because there were no publications before 1998; there was no significant trend overall from 1998-2014. The recent emergence of AGD in Norway and Scotland may be expected to increase scientific interest in the disease. Although *G. salaris* remains a serious concern, its spread appears to have been contained, and at great economic and ecological cost, the pathogen has been eradicated from some infected river systems (Mo et al. 2008). This may be the reason the publication rate has declined.

4.2.2.4 *Fungal diseases*

No fungal diseases are amongst the high H-indices pathogens for salmon, which corroborates the low ranking results found in human and domestic animals (McIntyre et al., 2014a). The fungal and oomycetes diseases, which are most important in freshwater (van den Berg et al. 2013), have historically been controlled using malachite green and formaldehyde. However these substances are toxic; malachite green has been banned worldwide and formaldehyde may well soon be banned in the EU. In the absence of these control substances, fungal problems could increase in significance and this would require scientific work to investigate both the epidemiology of outbreaks and the development of new medicines and treatment practices. Losses of 10% in hatcheries are already commonplace with a £5M cost in Scotland alone (van den Berg et al. 2013). Emerging fungal problems may be a threat to future food security. Searches of data on existing publications or assessments of impacts cannot identify future problems unless they have at least partially emerged (Cox et al.2016).

4.3 Limitations of the H-index

The calculated H-indices relate to salmon production and disease presence at the national level. They reflect our understanding of the importance and history of different diseases. The H-index method is a useful indicator for ranking salmon diseases, however, it has limitations. For example, it is conservative, may be biased against countries with limited resources, and there is limited independent data to validate its use in prioritisation.

The H-index method is conservative in that diseases of historic importance may continue to have high scores, although they are no longer of significance in terms of their impact. Conversely, emerging diseases take a few years to establish a score. Research itself can be conservative, with work building on increasing details of the properties of pathogens beyond specific applications to applied problems. The TI search helps reduce this bias because papers on general properties of pathogens are less likely to be selected. Historical change has been analysed using H-indices calculated on publications for different time periods and also by analysis of publication rates for key diseases. These identify expected trends of a reduction in impact of bacterial diseases and an increase in impact for viral diseases and sea lice.

H-indices relate overall to national salmon production, but exceptions occur for areas that lack the resources to invest (Hunter 2009). This may have been particularly important for Chile which, as a developing country, has had less financial resources. This may lead to lower scores for diseases that are of more concern to Chile than other countries, notably SRS, reflecting an under-investment in relevant science. Recent increases in Chilean scientific publication, however, indicate that this problem of resources may be being addressed.

The most serious limitation with the use of the H-index method is the lack of measures to assess its validity as a rank of disease priority (Brooks et al. 2015). This is not the case with diseases of humans or terrestrial animals for which extensive work on assessing impacts to allow ranking (del Rio Vilas et al. 2013, Brooks et al. 2015) and measures such as DALY have allowed the H-index to be validated in these contexts (McIntyre et al. 2011). There is a need for such analytical methods to be developed more extensively for aquaculture and for data to be more widely available and internationally comparable. Here validation of the use of the H-index method for ranking salmon diseases and parasites is limited to consistency with expectation for relative impact, trends in publication rates with time, and an assessment of presence/absence with non-zero H-indices. This is, however, also the reason that this method is useful; there is no alternative measure that can be widely applied in a consistent manner between diseases and countries for diseases of aquatic animals.

Funding of activity on disease control requires that these diseases be prioritised (Brooks et al. 2015). For example analyses of activities that risk spreading disease may require a list of pathogens for which risks are calculated (Murray 2015). Assessment of the general susceptibility and vulnerability of the system to emergence of diseases requires a list of key pathogens to be identified, before scenarios with specific properties can be investigated, since effectiveness of controls depends on pathogen transmission properties (Werkman et al. 2011). Research budgets need to be targeted to diseases that are of national importance. The H-index method is a useful guide for targeting priorities for action on current disease problems, but novel emerging diseases require assessment based on basic epidemiological principles of aquatic disease transmission (Murray and Peeler 2005) to identify likely future priorities.

Acknowledgements: This work was funded by the Scottish Government under AQ0050 and BBSRC Tools and Resources Development Fund (award BB/K003798/1)

Alderman D., Hastings T. 1998. Antibiotic use in aquaculture: developments of antibiotic resistance – potential for consumer health risks Food Sci. Technol. 33, 139-155

Anon. 2015. Overview - Web of Science. Thomson Reuters. URL: http://wok.mimas.ac.uk/. Accessed: 6/5/15

Aunsmo A., Valle P.S., Sandberg M., Midtlying P.J., Bruheim T. 2010. Stochastic modelling of direct costs of pancreas disease (PD) in Norwegian farmed Atlantic salmon (*Salmo salar* L.) Prev. Vet. Med. 93, 233-241

Brooks V.J., Hernández-Jover, M., Cowled, B., Holyoake, P.K., Ward, M.P. 2014. Building a picture prioritisation of exotic diseases for the pig industry in Australia using multi-criteria decision analysis. Prev. Vet. Med. 113, 103-117

Brooks V.J., del Rio Vilas V.J., Ward M.P. 2015 Disease prioritisation: what is the state of the art? Epidemiol. Infect. 143, 2911-2922

Costello M.J. 2009a. The global economic cost of sea lice to the salmonid farming industry. J. Fish Dis. 32, 115-118

Costello M.J. 2009b. How sea lice from salmon farms may cause wild salmonid declines in Europe and North America and be a threat to wild fish elsewhere. Proc. Roy. Soc. B, 275, 3385-3394

Cox R., McIntyre K.M., Sanchez J., Setzkorn C., Baylis M., Revie C.W. 2016. Comparison of the h-index scores for pathogens identified as emerging hazards in North America. Transb. Emerg. Dis. 63, 79-91

Crawley M.J. 2013. The R Book. John Wiley & Sons Ltd, Chichester, UK

Del Rio Vilas V.J., Voller F., Montibeller G., Franco, L.A., Sribhashyam S., Watson E., Hartley M., Gibbens J.C. 2013 An integrated process and management tool for ranking multiple emerging threats to animal health. Prev. Vet. Med. 108, 94-102

FAO 2014. FAO Global Aquaculture Production Volume and Value Statistics Database Updated to 2012. Food and Agriculture Organisation of the United Nations, Rome ftp://ftp.fao.org/FI/STAT/Overviews/AquacultureStatistics2012.pdf

Fofana A., Baulcomb C. 2012. Counting the costs of farmed salmonids disease. J. Appl. Aquacult. 24, 118-136

Hall M., Soje J., Kilburn R., Maquire S., Murray A.G. 2014. Cost-effectiveness of alternative disease management policies for bacterial kidney disease in Atlantic salmon aquaculture. Aquaculture 434, 88-92

Håstein T, Gudding, R., Evensen O. 2005. Bacterial vaccines for fish – an update on the current situation worldwide. Dev. Biol. 121, 55-74

Hastings T., Olivier G., Cusack R., Bricknell I., Nylund A., Binde M., Munro P., Allan C. 1999. Infectious salmon anaemia. Bull. Eur. Assoc. Fish Pathol. 16, 286-288

Hunter P.R. 2009. Bibliometrics, research quality, and neglected tropical diseases. Lancet 373, 630-631

International Committee on Taxonomy of Viruses (ICTV) 2015. URL: http://www.ictvonline.org/. Accessed: 2/9/15.

Jones S., Beamish R. 2011. Salmon lice: an integrated approach to understanding parasite abundance and distribution. Wiley-Blackwell, Chichester, UK

Kilburn R., Murray A.G., Hall M., Bruno D.W., Cockerill D., Raynard R.S. 2012. Analysis of a company's production data to describe the epidemiology and persistence of pancreas disease in Atlantic salmon (*Salmon salar* L.) farms of Western Scotland. Aquaculture 368-369, 89-94

King D.A. 2004. The scientific impact of nations. Nature 430, 311-316

Mackie T.J. 1935. Final report of the furunculosis committee. His Majesty's Stationary Office, Edinburgh UK

Mardones F.O., Perez, A.M., Vades-Donoso P., Carpenter T.E. 2011. Farm-level reproduction number during an epidemic of infectious salmon anaemia virus in southern Chile in 2007-2009. Prev. Vet. Med. 102, 175-184

McIntyre K.M., Hawkes, I., Waret-Szkuta, A., Morand, S., Baylis, M. 2011. The H-index as a quantitative indicator of the relative impact of human diseases. PLOS One

6, e19558 DOI: 10.1371/journal.pone.0019558 http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0019558

McIntyre K.M., Setzkorn, C., Hepworth, P.J., Morand, S., Morse, A.P., Baylis, M. 2014a. A quantitative prioritisation of human and domestic animal pathogens in Europe PLOS One 9, e103529 DOI: 10.1371/journal.pone.0103529 http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0103529

McIntyre K.M., Setzkorn C., Wardeh M., Hepworth P.J., Radford A.D., Baylis M. 2014b Using open-access taxonomic and spatial information to create a comprehensive database for the study of mammalian and avian livestock and pet infections. Prev. Vet. Med. 116, 325-335

Metselaar M., Thompson K.D., Gratacap R.M.L., Kik M.J.L., LaPatra S.E., Llyod S.J., Call D.R., Smith P.D., Adams A. 2010. Association of red-mark syndrome with a Rickettsia-like organism and its connection with strawberry disease in the USA. J. Fish Dis. 33, 849-858

Metzger D.C., Elliot D.G., Wargo A., Park L.K., Purcell M.K. 2010. Pathological and immunological responses associated with differential survival of Chinook salmon following *Renibacterium salmoninarum* challenge. Dis. Aquat. Org. 90, 31-41

Middlemas, S.J., Fryer, R.J., Tulett D. and Armstrong J.D. 2012. Relationship between sea lice levels on sea trout and fish farm activity in western Scotland. Fish. Manag. Ecol. 20, 68-74

Mo T.A., Norheim K., Jansen P.A. 2008 The surveillance and control programme for *Gyrodactylus salaris* in Atlantic salmon and rainbow trout in Norway. In Brun E., Jordsmyr H.M., Hellberg H, Mørk T. (Eds) Surveillance and control programmes for terrestrial and aquatic animals in Norway. Annual Report 2007, Oslo National Veterinary Institute 2008 p145-148

Munday B.L., Carson J., Whittington R., Alexander J. 1992 Serological responses and immunity produced in salmonids by vaccination with Australian strains of *Vibrio anguillarum* Immunol. Cell Biol. 70, 391–395

Murray A.G. 2006. A model of the spread of infectious pancreatic necrosis virus in Scottish salmon farms 1996-2003. Ecol. Modell. 199, 64-72

Murray A.G. 2015. Does the use of salmon frames as bait for lobster/crab creels significantly increase the risk of disease in farmed salmon in Scotland? Prev. Vet. Med. 120, 357-366

Murray A.G., Peeler E.J. 2005. A framework for understanding the potential for emerging diseases in aquaculture Prev. Vet. Med. 67, 223-235

Murray A.G, Munro L.A., Wallace I.S., Berx B., Pendrey D., Fraser D., Raynard R.S. 2010. Epidemiological factors in the re-emergence and control of an outbreak of infectious salmon anaemia in the Shetland Islands, Scotland. Dis. Aquat. Org. 91, 189-200

Murray A.G., Munro L.A., Wallace I.S., Allan C.E.T., Peeler E.J., Thrush M.A. 2012. Epidemiology of *Renibacterium salmoninarum* in Scotland and the potential for compartmentalised management of salmon and trout farming areas. Aquaculture 324-325, 1-13

National Center for Biotechnology Information 2015. U.S. National Library of Medicine, Bethesda, Maryland, US. The NCBI Taxonomy database homepage. URL: http://www.ncbi.nlm.nih.gov/Taxonomy/. Accessed: 2/9/2015.

OIE 2015. Manual of diagnostic tests for aquatic animals 2015. World Organisation for Animal Health, Paris. http://www.oie.int/international-standard-setting/aquatic-manual/access-online/

Rozas M., Enriquez R. 2014. Piscirickettsiosis and *Piscirickettsia salmonis* in fish: a review. J. Fish Dis. 37, 163-188

Soares S., Turnbull J.F., Green D.M., Crumlish M., Murray A.G. 2011. A baseline method for benchmarking mortality losses in Atlantic salmon (*Salmo salar*) production. Aquaculture 314, 7-12

Tavornpanich S., Paul M., Viljugrein H., Abrial D., Jimenez D., Brun E. 2012. Risk map and spatial determinants of pancreas disease in the marine phase of Norwegian Atlantic salmon farming sites BMC Vet. Res. 8, 172 DOI:10.1186/1746-6148-8-172

van den Berg, A.H., McLaggan,D., Diéguez-Uribeondo J., van West P. 2013. The impacts of the water molds *Saprolegnia diclina* and *Saprolegnia parasitica* on natural ecosystems and the aquaculture industry. Fungal Biol. Rev. 27, 33-42

Venkatachalam L. 2004. The contingent valuation method: a review. Environ. Impact Assess. Rev. 24, 89-124

Wardeh M., Risley C., McIntyre M.K., Setzkorn C., Baylis M. 2015. Database of host-pathogen and related species interactions, and their global distribution. Sci. Data 2:150049

Werkman M., Green, D.M., Murray, A.G., Turnbull J.F. 2011. The effectiveness of fallowing strategies in disease control in salmon aquaculture assessed with an SIS model. Prev. Vet. Med. 98, 64-73

Figure 1 H-index scores for diseases or pathogens calculated using searches of paper titles (TI) versus using full text in papers (TS) (r^2 = 0.72). Diseases at or beyond the 95% confidence range for the regression line with a TI H-index score > 0 are identified on the figure (see Table 1 for more detail), as are the diseases lice and ISA, as they are further discussed.

Figure 2. Mean H-index scores for the top nine salmon pathogens (see Table 2) by country versus natural logarithm of salmon production for 2012. The solid line illustrates a log regression of H-indices averaged for the different pathogens against national production, calculated excluding Chile (2.72 × ln(production) – 5.28 with 95% confidence intervals; $r^2 = 0.721$).

Figure 3. Number of publications describing diseases or pathogens of the Atlantic salmon (*Salmo salar*) by year for title (TI) searches. Publication rates for papers alluding to bacterial diseases (3a) showed significant declines for vibrioisis ($-0.147\ y^{-1}$, p =0.005) and furunculosis ($-0.318\ y^{-1}$, p =0.003). Viral diseases (3b) showed significant increases: ISA 0.21 y^{-1}, p =0.009; IPN 0.20 y^{-1}, p =0.001; PD 0.19 y^{-1}, p =0.015. Publications increased significantly for sea lice (0.70 y^{-1}, p =0.0005 and AGD (0.028 p = 0.001) but for *G. salaris* they had peaked and showed no significant changes (3c).

Figure 4. Number of publications from 2000-2014 by year describing salmon, published in Chile or Scotland); this demonstrates stable numbers in Scotland and a rapid increase in publication rates in Chile.

Table 1. Diseases and pathogens included in analyses, incorporating synonyms for both disease and pathogen

AGD: ("Amoebic Gill Disease" OR "*Neoparamoeba perurans" OR "Paramoeba perurans" OR "Paramoeba pemaquidensis")*

BKD*: ("Bacterial Kidney Disease" OR "Renibacterium salmoninarum")*

CMS*: ("Cardiomyopathy syndrome" OR "Piscine myocarditis virus")*

D. spath.: ("*Diplostomum spathaceum")*

EHN*: ("*Epizootic haematopoietic necrosis" OR *"*Epizootic hematopoietic necrosis")

ERM*: ("Enteric redmouth" OR "Yersinia ruckeri")*

Epitheliocystis: (Epitheliocystis OR "*Candidatus Piscichlamydia salmonis" OR" Candidatus Clavochlamydia salmonicola")*

Furunculosis*: ((Furunculosis NOT atypical) OR ("Aeromonas salmonicida" NOT atypical))*

G.salaris*: (Gyrodactylosis OR "Gyrodactylus salaris")*

HSMI: *("*Heart and skeletal muscle inflammation" OR "Piscine reovirus")

IHN*: ("*Infectious haematopoietic necrosis" OR *"*Infectious hematopoietic necrosis")

ISA: *("infectious salmon anaemia" OR "infectious salmon anemia")*

IPN*: ("*Infectious pancreatic necrosis")

PD: ("Pancreas disease" OR "sleeping disease" or "salmonid alphavirus")

PKD: ("Proliferative Kidney Disease" OR "*Tetracapsuloides bryosalmonae")*

RMS; ("Red mark syndrome")

Red vent: ("red vent" OR "Anisakis simplex")

Saprolegnia: (Saprolegnia)

Sarcocystis: (sarcocystis)

Sea lice: ("sea lice" OR "Lepeophtheirus salmonis" OR "Caligus elongatus" OR "Caligus rogercresseyi" OR "Caligus clemensi")

SRS: (Piscirickettsiosis OR "Salmon rickettsial syndrome" OR "*Piscirickettsia salmonis")*

Vibriosis: ("Vibriosis" OR "Listonella anguillarum" OR "Listonella anguillara" OR "Achromobacter ichthyodermis" OR "Pseudomonas ichthyodermis" OR "Vibrio ichthyodermis" OR "Vibrio piscium" OR "Vibrio anguillarum" OR "Vibrio salmonicida")

VHS: ("Viral haemorrhagic septicaemia" OR "Viral hemorrhagic septicemia")

Table 2. H-index scores for diseases or pathogens, calculated using title (TI = **bold**) or full text (TS = normal text), including synonyms or shortened names. For further details of diseases and pathogens see Table 1.

Disease or pathogen	Synonym/short name	TI	TS
Amoebic Gill Disease	AGD	***23***	28
Bacterial Kidney Disease	BKD	***12***	29
CardioMyopathy Syndrome	CMS	***9***	13
Diplostomum spathaceum	D. spath	***0***	9
Epizootic haematopoietic necrosis	EHN	***0***	5
Enteric Redmouth	ERM	***5***	19
Epitheliocystis	Epitheliocystis	***5***	15
Furunculosis	Furunc	***32***	67
Gyrodactylus salaris	G.salaris	***22***	36
Heart and Skeletal Muscle Inflammation	HSMI	***10***	12
Infectious haematopoietic necrosis	IHN	***8***	33
Infectious Salmon Anaemia	ISA	***28***	43
Infectious Pancreatic Necrosis	IPN	***24***	35
Pancreas Disease	PD	***22***	29
Proliferative Kidney Disease	PKD	***3***	20
Red Mark Syndrome	RMS	***0***	0
Red vent	Red Vent	***3***	12
Saprolegnia	Saprolegnia	***4***	12
Sarcocystis	Sarcocystis	***0***	1
Sea lice	Sea lice	***41***	54
Salmon rickettsial syndrome	SRS	***8***	22
Vibriosis	Vibriosis	***17***	53
Viral haemorrhagic septicaemia	VHS	***2***	28

Table 3. H-index scores calculated using paper titles (TI) for the top nine diseases and pathogens of Atlantic salmon, using bibliometric searches excluding any restriction to a specific country (All) or including publications for individual countries (Scotland-Scot, Norway including Faeroe Islands-Nor, Chile, Canada-Can, Denmark-Den, Australia-Aus, USA, Ireland-Ire).

Synonym/short name for disease/ pathogen	**All**	**Scot**	**Nor**	**Chile**	**Can**	**Den**	**Aus**	**USA**	**Ire**
BKD	***12***	3	2	0	7	0	0	2	0
Vibriosis	***17***	2	14	1	0	0	2	2	0
G. salaris	***22***	3	19	0	0	6	0	1	0
AGD	***23***	1	3	1	5	0	22	1	2
PD	***22***	11	12	0	0	2	0	0	18
IPN	***24***	16	16	0	1	1	0	1	1
ISA	***28***	9	25	3	14	4	0	5	1
Furunc	***32***	13	28	0	10	3	0	9	6
Sea lice	***41***	30	25	6	24	1	2	16	11