Preventive Veterinary Medicine 107 (2012) 280–285

Contents lists available at SciVerse ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Short communication

Prioritizing live bird markets at risk of avian influenza H5N1 virus contamination for intervention: A simple tool for low resource settings

Gina Samaan[a,*], Risa Indriani[b], Luis Roman Carrasco[c], Kamalini Lokuge[a], Alex R. Cook[d,e,f], Paul M. Kelly[a,g], Rma Adjid[b]

[a] National Centre for Epidemiology and Population Health, College of Medicine, Biology & Environment, The Australian National University, Canberra, ACT 0200, Australia
[b] Indonesian Research Center for Veterinary Science, Agency for Agricultural Research and Development, Ministry of Agriculture, Jalan RE Martadinata 30, Bogor, Indonesia
[c] Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore
[d] Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117597, Singapore
[e] Department of Statistics and Applied Probability, Faculty of Science, National University Singapore, Singapore 117546, Singapore
[f] Program in Health Services and Systems Research, Duke-NUS Graduate Medical School Singapore, 8 College Road, Singapore 169857, Singapore
[g] Population Health Division, ACT Government Health Directorate, GPO Box 825, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 14 December 2010
Received in revised form 24 May 2012
Accepted 31 May 2012

Keywords:
Avian influenza
H5N1
Live bird markets
Decision tree
Indonesia
Screening

ABSTRACT

Live bird markets (LBMs) are at risk of contamination with the avian influenza H5N1 virus. There are a number of methods for prioritizing LBMs for intervention to curb the risk of contamination. Selecting a method depends on diagnostic objective and disease prevalence. In a low resource setting, options for prioritization are constricted by the cost of and resources available for tool development and administration, as well as the resources available for intervention. In this setting, tools can be developed using previously collected data on risk factors for contamination, and translated into prediction equations, including decision trees (DTs). DTs are a graphical type of classifier that combine simple questions about the data in an intuitive way. DTs can be used to develop tools tailored to different diagnostic objectives. To demonstrate the utility of this method, risk factor data arising from a previous cross-sectional study in 83 LBMs in Indonesia were used to construct DTs. A DT with high specificity was selected for the initial stage of an LBM intervention campaign in which authorities aim to focus intervention resources on a small set of LBMs that are at near-certain risk of contamination. Another DT with high sensitivity was selected for later stages in an intervention campaign in which authorities aim to detect and prioritize all LBMs with the risk factors for virus contamination. The best specific DT achieved specificity of 77% and the best sensitive DT achieved sensitivity of 90%. The specific DT had two variables: the size of the duck population in the LBM and the human population density in the LBM's district. The sensitive DT had three variables: LBM location, whether solid waste was removed from the LBM daily and whether the LBM was zoned to separate the bird holding, slaughtering and sale areas. High specificity or sensitivity will be preferred by authorities depending on the stage of the intervention campaign. The study demonstrates that simple tools utilizing DTs can be developed to prioritize LBMs for intervention to control H5N1-virus. DT tools are simple to apply, suitable for low-resource settings and can be tailored to the particular needs and stage of the disease control program.

* Corresponding author. Tel.: +62 813 1753 3978.
E-mail address: ginasamaan@yahoo.com (G. Samaan).

0167-5877/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.prevetmed.2012.05.017

1. Background

The avian influenza A H5N1 virus is of global public health concern due to its high pathogenicity in birds, its current zoonotic capability and its pandemic potential (Briand and Fukuda, 2009). In virus endemic areas, live bird markets (LBMs) are vulnerable to contamination since bird populations coming into the LBM are dynamic and infected flocks may enter at any time. This increases the risk of virus transmission both to humans and animals in the LBM, and it increases the risk of propagating virus back into farms through the sale of infected live birds (Kung et al., 2003; Wang et al., 2006).

Previous research has shown that risk factors for H5N1-virus contamination in the LBM environment such as surfaces, floors and utensils include slaughtering birds in the LBM, lack of zoning in the poultry workflow and insufficient waste management (Bulaga et al., 2003; Garber et al., 2007; Indriani et al., 2010). In high resource settings, LBMs are managed through the enhanced application and monitoring of practices for good general hygiene and disease control (Lu, 1970; Mullaney, 2003; Trock et al., 2008). In low resource settings, there is limited capacity for hygiene and authorities do not have sufficient resources to intervene in all LBMs. Thus, the key question is 'how do authorities prioritize LBMs to invest their limited resources for disease control?'

Based on principles of screening and diagnostic testing for disease control (Wilson and Jungner, 1968), there are a number of methods for prioritizing LBMs for intervention (Table 1). Selecting a method depends on its fitness-for-purpose, including cost, sensitivity, specificity, speed, complexity as well as human and hardware resource requirements. In a low resource setting, options for prioritization are constricted by the cost of and resources available for tool development and administration, as well as the resources available for intervention. This decreases the feasibility of using laboratory-based tools and network analyses in LBMs as they are expensive to develop and administer, and mandate laboratory or statistical expertise (Table 1). Thus, other options need to be explored.

A number of low resource countries affected by the H5N1-virus have conducted cross-sectional surveys in LBMs to assess virus prevalence in birds and the LBM environment (Abdelwhab et al., 2010; Indriani et al., 2010; Jiang et al., 2010; Negovetich et al., 2011). Data from such studies can be used to develop tools to prioritize LBMs for intervention. Risk factor data can be translated into prioritization tools using prediction equations including classifiers such as decision tree (DTs). DTs categorize LBMs into groups, where those with the risk factors for contamination are deemed priority. These LBMs can then be targeted for public/veterinary health action, maximizing utilization of public health resources in low resource settings (World Health Organization, 2006).

DTs are quick and relatively simple to administer and interpret (Table 1). While DTs constitute one way of presenting and communicating results derived from prediction equations, well-established alternatives include logistic regression models that can be presented as predictive probabilities or as odds-ratios. The utility of any of these tools to veterinary and public health practitioners depends on the epidemiological considerations and the diagnostic criteria established. Using a previous cross-sectional study conducted in 83 LBMs in Indonesia as a case study, we explored DT options in tools for prioritizing LBMs for interventions based on different epidemiological considerations.

2. Methods

2.1. Data and problem formulation

The tools were developed for national authorities intervening in all LBMs in the three provinces reported in the Indriani et al. study: Banten, Jakarta and West Java. The Indriani et al. study assessed environmental contamination and risk factors for contamination in 83 LBMs randomly selected from the 300 LBM population. The Indriani et al. study found that at least one environmental site in 39 of the 83 LBMs (47%) tested were contaminated with the H5N1-virus. The ten risk factors identified from the univariate analysis were used to develop candidate DTs (Indriani et al., 2010). We also considered two variables known as risk factors for H5N1-virus spread in the three target provinces: density of farmed birds (chickens and duck) and human population density at district level (Loth et al., 2011). Details of the variables considered in the model development can be seen in Supplementary File 1.

2.2. Epidemiological considerations

Diagnostic objective and disease prevalence guided DT design.

2.2.1. Diagnostic objective

The stage of disease control and intervention resources available guide the decision to optimize diagnostic sensitivity or specificity. At the beginning of an intervention campaign, where authorities aim to reduce the overall level of contamination and virus circulation in LBMs, high specificity will limit the number of LBMs deemed priority and ensure that limited resources for intervention are allocated most efficiently. That is, if resources are available to intervene in only a proportion of LBMs, this tool allows us to maximize the number of infected LBMs in the sub-group receiving interventions. Even though this approach may yield low sensitivity, it is operationally more feasible based on constricted resources available for intervention. Further, unlike diagnostic tests in which outcomes may be catastrophic for the patient or for the unit receiving the intervention, the implications of a low sensitivity here are not dire for individual LBMs. In an intervention campaign, when levels of virus are reduced such that eradiation is a feasible objective, high sensitivity will ensure that all LBMs with the risk factors will be prioritized for intervention. Even though a tool with high sensitivity may risk low specificity, this may be acceptable to authorities since eradication is in near sight and the absolute number of LBMs deemed priority will be small.

Table 1
Characteristics of tools that can trigger disease control interventions for H5N1-virus in LBMs.

Tool target and method	Utility in program	Tool development and administration				Generalizability of method
		Development cost	Administration cost	Speed	Complexity	
Target: Detect LBMs contaminated with virus *Method:* laboratory testing (e.g. VI or PCR) of samples collected from birds or surfaces	Ongoing	Moderate	High	Moderate	High	Moderate – laboratory capacity needed
Target: Prioritize LBMs based on connectivity and potential to spread virus outside LBM *Method:* network analysis	Initial stages	Moderate	High	Slow	High	Low – expertise in method needed
Target: Prioritize LBMs based on risk factors for contamination *Method:* Prediction equations including DT	(a) Initial stage (if specific) (b) Latter stage (if sensitive)	Moderate	Low	Fast	Low	Moderate – depends on availability of risk factor data and knowledge in statistical technique

VI, virus isolation; LBM, live bird market; PCR, polymerase chain reaction; DT, decision trees; LBM, live bird market.

2.2.2. Disease prevalence

Prevalence, also known as the prior probability of disease, impacts the predictive values of a tool. In populations of LBMs where contamination is common, the tool's positive predictive value (PPV) will be higher than in a population of LBMs where the H5N1-virus contamination is rare. The converse is true for negative predictive value (NPV). In areas endemic for H5N1-virus, the prevalence is expected to be high since all LBMs are at risk of contamination. As was seen in the Indriani et al. study, 47% of LBMs were contaminated with the virus. Another cross-sectional study in LBMs in Egypt found H5N1-virus prevalence ranged from 27% to 41% depending on season (Abdelwhab et al., 2010). Ultimately, users of the tool need to be mindful that the predictive values observed in one study do not apply universally, and thus, require that the DTs underlying the tool be redeveloped in other settings and periodically reviewed in response to changes in prevalence.

2.3. Statistical methodology

We generated candidate DT models based on four groups of variables: (a) ten risk factors on univariate analysis from the Indriani et al. study, (b) four independent risk factors from the Indriani et al. study, (c) ten risk factors on univariate analysis from the Indriani et al. study in addition to the two risk factors identified in the Loth et al study, and (d) the four independent risk factors from the Indriani et al. study in addition to the two risk factors identified in the Loth et al study. For each candidate model, we fitted the model to all three provinces as well as to each province alone. We also constructed and compared pruned and unpruned DTs for all candidate models. From the total 45 candidate models generated, we selected those that provided the highest specificity or sensitivity, along with the highest discrimination ability based on the area under the ROC curve (AUC). We used classification tree models to construct the DTs. This method is commonly used to translate risk factor data into a diagnostic tool by combining simple if–then questions about the data (Kingsford and Salzberg, 2008). As no independent validation set was available, leave-one-out cross-validation was employed to estimate the robustness of the DTs and their predictive power (Ripley, 1996). Analyses were done using R statistical environment. Further technical details of the DT method can be seen in Supplementary File 2.

The dependent variable "true contamination" was defined as at least one of the 27 environmental sites in the LBM being positive by real time reverse transcription–polymerase chain reaction (RRT–PCR) based on the laboratory findings conducted in the Indriani et al. study. RRT–PCR methods are known to present high, but not 100% sensitivity and specificity in the detection of influenza A samples (Spackman et al., 2003). We performed two different fits of the trees to the data: (i) we assumed 100% specificity and sensitivity of the RRT–PCR and assumed that the observed data represented true positives and negatives; (ii) we estimated a sensitivity of 0.66 and specificity of 0.95 for the RRT–PCR by comparing with the current standard for detection of influenza virus (virus isolation in embryonated chicken eggs) which was used as a benchmark (Spackman et al., 2003). The original data set was modified following a stochastic process that mimicked the sensitivity and specificity of the RRT–PCR when detecting H5N1-virus in the LBMs samples. This was done by modifying the original dataset by introducing false negatives and positives with a probability corresponding to the sensitivity and specificity of the RRT–PCR. For instance, if a given sample was positive, a random number from zero to one was generated, if the random number is above the sensitivity of the RRT–PCR, the sample is changed to negative (that is, a false negative is generated) and if it is below it remains as positive.

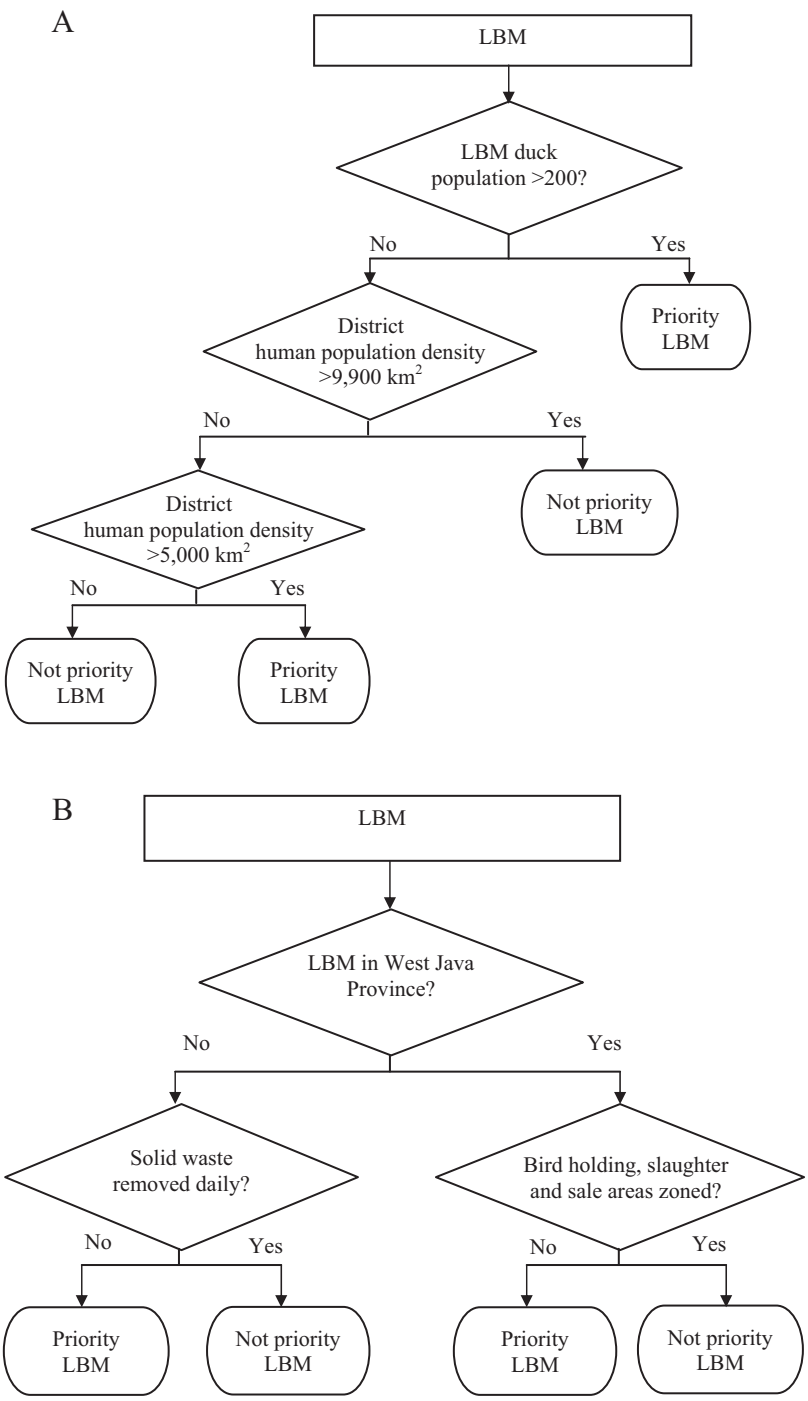

Fig. 1. Decision trees optimizing diagnostic specificity (1A) and sensitivity (1B) for determining priority LBMs for intervention.

The introduction of such noise allowed us to test the sensitivity of our results to the noise that is expected to be introduced by the RRT–PCR test. The models were fitted to 1000 different modifications of the original data using the described procedure and the mean and standard deviation of the sensitivity and specificity of the models were stored.

3. Results

The specific DT had two variables: the size of the duck population in the LBM and the human population density in the LBM's district (Fig. 1A). This DT had three splits, achieved 77% specificity, 36% sensitivity and had 48% AUC. When considering RRT–PCR as an imperfect

detection system, specificity decreased to an average of 65% (standard deviation se of 3%) and sensitivity remained at an average of 36% ($se = 5\%$). The sensitive DT had three variables: LBM location, whether solid waste was removed from the LBM daily and whether the LBM was zoned to separate the bird holding, slaughtering and sale areas (Fig. 1B). This DT achieved 90% sensitivity, 43% specificity and had 57% AUC. When considering RRT–PCR as an imperfect detection system, sensitivity and specificity decreased to 87% ($se = 5\%$) and 36% ($se = 3\%$) respectively.

The structure of the variables in the models reflects variable interaction; the first split corresponding to the variable explaining the greatest amount of the variance, and subsequent splits explaining successively less. However, the DTs do not suggest an interaction (in the sense of regression models) between the variables. Analyses using generalized linear models yielded similar findings as the DTs presented; the same variable combinations for the specific and sensitive diagnostic objectives and similar diagnostic performance (data not shown).

4. Discussion

The study demonstrated how simple tools to prioritize LBMs for intervention can be developed using risk factor data arising from cross sectional research. The tools can be tailored for different epidemiological considerations, including target population and resources available for intervention. Since risk factor data are required to develop tools using DTs, the cost of tool development may be considerable. However, low resource countries affected by H5N1-virus have conducted the types of studies that provide the data required to develop these tools. Also, once the tools have been developed, they have low administration cost and can be applied by a variety of authorities including veterinarians, public health officers, sanitarians and market authorities. However, the DTs need to be tested and refined according to new data and varying field conditions.

Important aspects in deciding on diagnostic tests are fitness-for-purpose (Swayne, 2008). A tool with high specificity is useful in the early stages of an intervention campaign, whilst a tool with high sensitivity is useful towards the end to detect any remaining LBMs that have risk factors for contamination. However, neither tool is suitable for routine surveillance or to certify LBMs free of live virus.

The diagnostic performance of the tools in this study involves tradeoffs. In the target LBM population, the specific tool would miss many contaminated LBMs since sensitivity was low, and the sensitive tool would deem many contamination-free LBMs priority since specificity was low. Even though tools with high specificity or sensitivity increased the error rate for the counterpart, this may nevertheless be acceptable to authorities depending on the objective and stage of the intervention campaign.

There are two main limitations in the current study. Firstly, due to resource limitations for the study, there was no independent validation set. To overcome this, we used leave-one-out cross-validation which provides a good measure of out of sample predictive performance. The second limitation is that the tools generated here have limited generalisability to external LBM populations. The tools were developed for a targeted LBM population and the variables considered included those specific to the three provinces, such as human population density and LBM location. The provinces for which these tools were developed have the most human H5N1 cases and reports of outbreaks in birds nationally (Loth et al., 2011), which may warrant this region's prioritization of H5N1 disease control in LBMs. For other settings, we believe the study highlighted the epidemiological considerations and statistical approach to developing tools by translating risk factor data into prediction rules that can be applied in the field. It is also important to note that the dependent variable was based on RRT–PCR positivity. Since RRT–PCR detects fragments of the virus, positivity does not directly correlate with current presence of viable virus. Positive RRT–PCR findings should nevertheless spur public health action in LBMs as the results indicate recent historical contamination.

In conclusion, we have shown that a simple tool utilizing DT can be developed to prioritize LBMs for intervention to control H5N1-virus. Even though a variety of methods exist to translate risk factor data into diagnostic tools, we focused on DTs as they are simple to apply, suitable for low-resource settings and can be tailored to the particular needs and stage of the disease control program.

Acknowledgements

Paul Kelly's salary is partly funded by Australia's National Health & Medical Research Council. Luis Roman Carrasco and Alex R Cook are grateful for research funding from the research grant NMRC/H1N1R/005/2009.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.prevetmed.2012.05.017.

References

Abdelwhab, E.M., Selim, A.A., Arafa, A., Galal, S., Kilany, W.H., Hassan, M.K., Aly, M.M., Hafez, M.H., 2010. Circulation of avian influenza H5N1 in live bird markets in Egypt. Avian Dis. 54, 911–914.

Briand, S., Fukuda, K., 2009. Avian influenza A (H5N1) virus and 2 fundamental questions. J. Infect. Dis. 199, 1717–1719.

Bulaga, L.L., Garber, L., Senne, D.A., Myers, T.J., Good, R., Wainwright, S., Trock, S., Suarez, D.L., 2003. Epidemiologic and surveillance studies on avian influenza in live-bird markets in New York and New Jersey, 2001. Avian Dis. 47, 996–1001.

Garber, L., Voelker, L., Hill, G., Rodriguez, J., 2007. Description of live poultry markets in the United States and factors associated with repeated presence of H5/H7 low-pathogenicity avian influenza virus. Avian Dis. 51, 417–420.

Indriani, R., Samaan, G., Gultom, A., Loth, L., Indryani, S., Adjid, R., Putu Indi Dharmayanti, N.L., Weaver, J., Mumford, E., Lokuge, K., Kelly, P.M., Darminto, 2010. Environmental sampling for avian influenza virus A (H5N1) in live-bird markets, Indonesia. Emerg. Infect. Dis. 16, 1889–1895.

Jiang, W.-M., Liu, S., Chen, J., Hou, G.-Y., Li, J.-P., Cao, Y.-F., Zhuang, Q.-Y., Li, Y., Huang, B.-X., Chen, J.-M., 2010. Molecular epidemiological surveys of H5 subtype highly pathogenic avian influenza viruses in poultry in China during 2007–2009. J. Gen. Virol. 91, 2491–2496.

Kingsford, C., Salzberg, S.L., 2008. What are decision trees? Nat. Biotechnol. 26, 1011–1013.

Kung, N.Y., Guan, Y., Perkins, N.R., Bissett, L., Ellis, T., Sims, L., Morris, R.S., Shortridge, K.F., Peiris, J.S., 2003. The impact of a monthly rest day on

avian influenza virus isolation rates in retail live poultry markets in Hong Kong. Avian Dis. 47, 1037–1041.
Loth, L., Gilbert, M., Wu, J., Czarnecki, C., Hidayat, M., Xiao, X., 2011. Identifying risk factors of highly pathogenic avian influenza (H5N1 subtype) in Indonesia. Prev. Vet. Med. 102, 50–58.
Lu, F.C., 1970. The joint FAO–WHO food standards programme and the codex alimentarius. WHO Chron. 24, 198–205.
Mullaney, R., 2003. Live-bird market closure activities in the northeastern United States. Avian Dis. 47, 1096–1098.
Negovetich, N.J., Feeroz, M.M., Jones-Engel, L., Walker, D., Alam, S.M.R., Hasan, K., Seiler, P., Ferguson, A., Friedman, K., Barman, S., Franks, J., Turner, J., Krauss, S., Webby, R.J., Webster, R.G., 2011. Live bird markets of Bangladesh: H9N2 viruses and the near absence of highly pathogenic H5N1 influenza. PLoS One 6, e19311.
Ripley, B.D., 1996. Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge.
Spackman, E., Senne, D.A., Bulaga, L.L., Myers, T.J., Perdue, M.L., Garber, L.P., Lohman, K., Daum, L.T., Suarez, D.L., 2003. Development of real-time RT-PCR for the detection of avian influenza virus. Avian Dis. 47, 1079–1082.
Swayne, D.E., 2008. Avian Influenza. Blackwell Pub, Ames, Iowa.
Trock, S.C., Gaeta, M., Gonzalez, A., Pederson, J.C., Senne, D.A., 2008. Evaluation of routine depopulation, cleaning, and disinfection procedures in the live bird markets, New York. Avian Dis. 52, 160–162.
Wang, M., Di, B., Zhou, D.H., Zheng, B.J., Jing, H., Lin, Y.P., Liu, Y.F., Wu, X.W., Qin, P.Z., Wang, Y.L., Jian, L.Y., Li, X.Z., Xu, J.X., Lu, E.J., Li, T.G., Xu, J., 2006. Food markets with live birds as source of avian influenza. Emerg. Infect. Dis. 12, 1773–1775.
Wilson, J.M.G., Jungner, G., 1968. Principles and Practice of Screening for Disease. World Health Oragnization.
World Health Organization, 2006. In: WHO Press (Ed.), A Guide to Healthy Food Markets. World Health Organization.