



**HAL**  
open science

## Use of health data for research

Marie-Hélène Pinard-van Der Laan

► **To cite this version:**

Marie-Hélène Pinard-van Der Laan. Use of health data for research. ICAR Conference - Challenges and benefits of health data recording in the context of food chain quality, management and breeding, May 2013, Aarhus, Denmark. pp.63-71. hal-01193945

**HAL Id: hal-01193945**

**<https://hal.science/hal-01193945>**

Submitted on 3 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

---

## Use of health data for research

*M.-H. Pinard-van der Laan*

*UMR GABI, Animal Genetics Division, INRA, 78352 Jouy en Josas Cedex, France*

Animal health issues are of increasing importance to all animal breeding sectors, by raising health and welfare issues and causing major production costs, but also to citizens, by affecting possibly their own health and lifestyle choices. So far, traditional selective breeding approaches have been applied successfully to enhance production and reproduction traits; but the inclusion of animal health related traits are scarcely considered because of a clear lack of easy measurable and relevant health phenotypes and associated genetic markers which could be integrated in running breeding programs. Though, there have been numerous studies in the past showing evidence of genetic variability of responses to various diseases of economic interest. New opportunities have been arising thanks to major advances in animal genomics and related technologies. Most research strategies are now developed, combining structural, population and functional genomics approaches. The objectives of this research studies using health data are two-fold: 1-identification of genes, gene products and regulatory networks involved in host pathogen interactions which could be used in selection and 2- better understanding their functions and the underlying mechanisms. Applications of such research using health data may cover areas like vaccine development, improved diagnosis and treatment, epidemiology, in addition to breeding. Likely, integrated strategies using several approaches will be more successful to combat diseases. Thus, research and development aiming at understanding health related mechanisms and implementing integrated strategies for improved health of livestock requires large use of recorded health data, collaborative multidisciplinary programs and efficient technology transfer between research and industry. Improving recording systems of health data would benefit both research and industry in an ultimate goal of breeding more "robust" animals for a sustainable production.

*Keywords: Animal disease, animal health, genomics, health data, phenotypes, recording.*

In the context of intensification and specialization of all livestock sectors, animal welfare and animal health are given an increased emphasis in current breeding objectives. Indeed, animal health issues are raising health and welfare problems but also causing major production costs, but also to citizens, by affecting possibly their own health and lifestyle choices. So far, traditional selective breeding approaches have been applied successfully to enhance production and reproduction

---

---

### Abstract

---

---

---

---

### Introduction

---

---

traits; but the inclusion of animal health related traits are scarcely considered because of a clear lack of easy measurable and relevant health phenotypes and associated genetic markers which could be integrated in running breeding programs. To gather all this necessary information, acquisition of relevant health data is crucial and can be organized and exploited at different complementary levels, as presented in this paper.

---

---

**Need of health data for an integrated approach of disease control**

---

---

Clearly, animal health issues may be better tackled in animal breeding by developing integrated approaches than by applying single disease control strategies. Integrated animal health management involves both prevention and cure of diseases, potentially important decisions on the animals, like selection, culling, and developing vaccination programmes. It concerns also the pathogen and its environment when applying chemotherapy or biosecurity measures.

Integrated disease control therefore comprises important steps with potentially major effects for the animal populations and their environments. Thus, it needs a prior critical evaluation of the disease targets and the optimal control means, requiring necessarily health data in quality and quantity.

---

---

**Need of health data for validating disease resistance phenotypes**

---

---

Among most promising approaches for prevention and control of animal diseases are - breeding for disease resistance and - development of novel vaccines. Developing breeding for disease resistance strategies for any given disease requires jointly identifying genetic markers and underlying mechanisms related to the disease resistance phenotypes. And there is, to date, a clear lack of disease phenotypes. This is a problem, well known as the "phenotypic gap", but which appears as an even more important bottleneck for health related traits as it is often more difficult to find easy to measure and relevant health traits.

An example of a classical approach to identify the desirable set of disease resistance traits is given. This example applies to a parasitic disease in poultry, but can be transposed to any other disease and animal species. Implementation of selection for disease resistance to coccidiosis is hampered by a lack of easily measurable, repeatable and relevant disease phenotypes. A two-fold implementation has been processed: from large to small set of phenotypes, and from experimental to commercial lines (Figure 1; Hamzic *et al.*, 2014).

Firstly, a large panel of disease phenotypes has been measured and analysed (repeatability, variability, correlations...), on experimental chicken lines showing large individual variability for the response to coccidiosis. Secondly, most pertinent measures have been validated, in pilot study, then in large scale challenge of commercial animals. Ideally, this small set of validated phenotypes, which can be automated, will be sufficient to exploit the genetic variability for further selection.

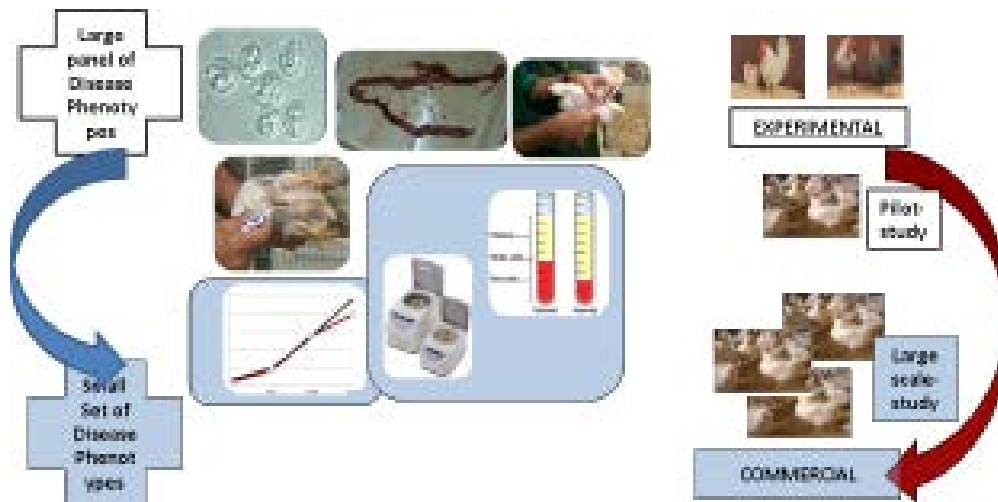


Figure 1. A classical approach to validate disease resistance phenotypes, from a large panel to a smaller set, from experimental to commercial.

Prior starting measuring a number of health related phenotypes to initiate a breeding for disease resistance strategy, a critical evaluation of the diseases is needed. A pertinent choice will look first at a "genetic variation" score, which includes 1- a well documented genetic variability for the diseases of interest and 2- the availability of genomic tools to measure and further exploit this variability. This latter criterion is likely to evolve fast as genomic tools become more sophisticated for each species and more information becomes available. But to prioritize diseases, one should consider also other criteria, without which practical interest of developing a breeding strategy is minimal, such as industry concern, economic impact, public concern, threat to food safety or zoonotic potential, impact on animal welfare..., making up a "disease score". The "genetic variation" score and the "disease" score" may be combined to establish a list of "priority diseases" (Figure 2; Davies *et al.*, 2009). This

**Need of health data for breeding for disease resistance. For which diseases?**

Disease	Disease score (Industry-eco-public-welfare-zoono.)	Genetic Variation	Priority
Mastitis	++	+++	↓
Bovine leukemia	+	++	
Gastrointestinal parasites	+	++	
Paratuberculosis	++	+	
Bovine TB	[*]	+	
Bacterial pneumonia	+	+	
E. coli	+++		
FMD	+++		
Brucellosis	+++		
IBR	[*]	+	
Salmonella spp	++		
BVD	++		

Figure 2. Priority list of infectious cattle diseases, with disease scores taking into account industry and public concern, economic and zoonotic potential, animal welfare impacts. (Simplified from Davies *et al.*, 2009).

ranking of diseases needs of course to be updated regularly within our changing environments and adapted to different international contexts if needed. This approach can very well be used to rank diseases from other perspectives than breeding, like in terms of disease control strategies.

**Identification of genetic markers and underlying mechanisms of health traits**

Once 1- disease resistance phenotypes have been identified and validated, these well characterized phenotypes are essential to identify 2- genetic markers and 3-underlying mechanisms, in a process of developing breeding for disease resistance strategies (Figure 3).

Though efforts in genome sequencing and SNP discovery are still needed, denser chips are now developed, allowing, through whole-genome association analysis, the identification of genetic markers associated with disease traits (e.g., identification of SNP associated with resistance to paratuberculosis in Holstein cattle, figure 4; Kirkpatrick *et al.*, 2011). Making use of high throughput technologies, individual genome sequences allow now the identification of the causal mutations associated with health traits and the generation of SNP which can be used in population-wide studies in order to identify animals with improved health related traits (Bai *et al.*, 2012). This new refine knowledge in the genetic control of health traits can bring additional understanding to a "black box" genomic selection process of livestock and can turn into novel tools for control and diagnostic of animal diseases (Figure 3).

The same advanced genome enabled technologies offer new opportunities to better understand regulation of the disease process, (e.g. immune responses, molecular dialogue between the host and the pathogen) by greater insight into transcriptomics and now sequencing of transcripts. Whereas QTL approaches aim at identifying mutations underlying genetic variations seen between hosts for the given health traits, transcriptomics approaches aim at identifying genes being transcribed in a particular tissue at a particular time of the disease process. Thus, both are essential

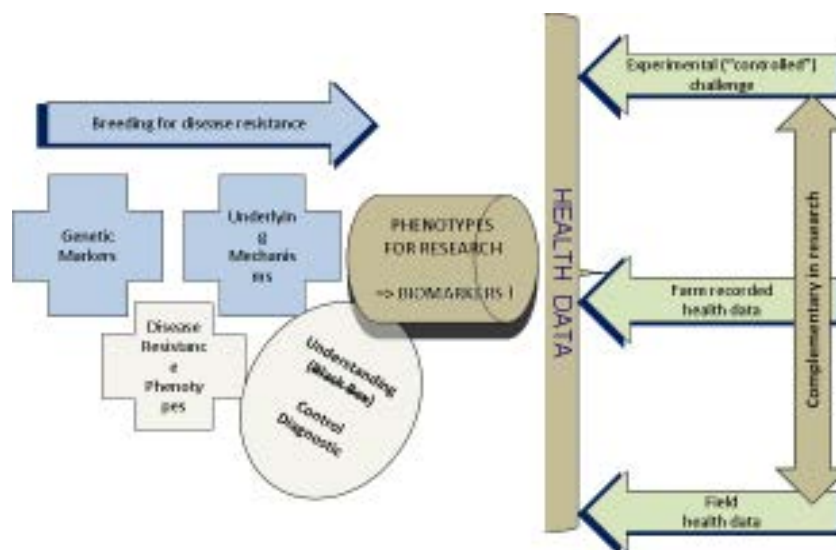


Figure 3. Use of health data, from complementary sources (experimental, farm, field), as novel phenotypes for research (in identifying genetic markers, underlying mechanism, disease resistance phenotypes).

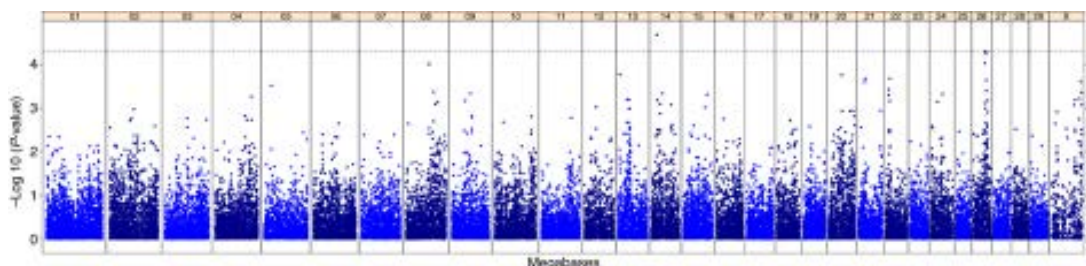


Figure 4. Whole-Genome association analysis of susceptibility to paratuberculosis in Holstein cattle (Kirkpatrick *et al.*, 2011).

and complementary approaches, giving tools to dissect, understand and utilize host genetic variation, using health data towards identification of genetic and functional biomarkers, the ultimate goal (Figure 3).

Because of the obvious need of developing integrated strategies to control diseases, not only complementary tools are needed (see above), but also complementary sources of large-scale health data are necessary, from 1- experimental "controlled" challenges, 2- farm and 3- field levels (Figure 3).

Experimental "controlled" challenges may allow the initial screening of health phenotypes (see § "Need of health data for validating disease resistance phenotypes"), as many measurements necessitate in the first steps to deliberately infect animals with control doses for instance. These health data will be highly complementary to measurements obtained in farms or accessed from the field by collecting veterinary records or history of naturally infected animals' episodes.

Use of large scale multi farm data will give the unique opportunity to have access to health data obtained in different environments, husbandry or management conditions, allowing to make use of differences to gain in standardization of the health phenotypes of interest. Ultimate goal, of prime interest for the breeders, is identifying health data as indicators of robustness, i.e., the ability of animals to cope with variable environments. Thus, having access to detailed health records and animal performances from a variety of environments, using these production traits as proxy for health is likely the only way to make a significant step towards pertinent definition of robustness and identification of usable markers for breeding. Maximizing animal resources can be then achieved through trusted collaboration between actors and facilitated dialogue between disciplines.

Field health data are often found to be complementary to farm recorded data, for example by making use of outbreaks happening in the field with health data recorded in cohorts in farms. Such complementary source of health data is applied on the PRRS (Porcine Reproductive and Respiratory) syndrome in pigs exemplifying a stakeholder cooperation with data used from cohorts of growing pigs from different breeding companies (Boddicker *et al.*, 2012). Epidemic field data can be also envisaged to be more efficiently used as done so far, for example to calibrate genomic predictions, as in the case of bovine tuberculosis (Bishop *et al.*, 2012). Finally, field

---



---

### Need of using complementary sources of health data

---



---



health data from animals exposed to recurrent disease threats or extreme environments, like in developing countries, offer often underexploited richness in terms of reservoir of phenotypes and genotypes.

**Best use of health data by collaborative research**

Significantly advances in knowledge in the field of animal health can best come from use of health data from different sources, but also from different species, combining multi-disciplinary expertise and eventually multi-national projects. Such an example of collaborative research aiming at sharing tools and optimizing animal resources and health data has been conducted in a large animal disease genomics network ([www.eadgene.info](http://www.eadgene.info)).

As an illustration, an original joint research programme on transcriptomic studies on mastitis in ruminants has been conducted, combining heterogeneous microarray data realized with different ruminant host species and infected with different mastitis-causing pathogens (Table 1; Genini *et al.*, 2011). For the first time, a meta-analysis combined experiments and time points to create the four main responses to mastitis infection. Such an analysis showed that it could fit well with the high heterogeneity of health data representing the different biological systems, being a proof of principle for studies combining health data from diverse sources. But most importantly, the results reinforced previous findings but also revealed several new

*Table 1. Combination of experiments and time points to create the 4 main responses to mastitis infection (from Genini et al., 2011).*

Experiment #	Time after infection							
	2 h	3 h	6 h	8 h	12 h	24 h	36 h-72 h	72 h
1A: <i>E. coli</i> in cattle ( <i>in vivo</i> )			{1}		{2}	{3}		
1B: <i>S. aureus</i> in cattle ( <i>in vivo</i> )			I, II, IV		I, IV	I, III, IV		
1C: <i>S. aureus</i> in cattle ( <i>in vivo</i> )			{4}		{5}	{6}		
2: <i>S. uberis</i> in cattle ( <i>in vivo</i> )			I, IV		I, IV	I, II, IV		
3: <i>S. aureus</i> in cattle macrophages ( <i>in vitro</i> )	{10}		{11}		I, II, IV			{8}
4: <i>S. aureus</i> in goat ( <i>in vivo</i> )	I, II, IV		I, III, IV				{9}	I, III, IV
5: <i>S. aureus</i> in sheep dendritic cells ( <i>in vitro</i> )		{14}		{15}	{12}	{13}		
6: <i>S. aureus</i> in goat ( <i>in vivo</i> )		I, II		I, III	I, II	I, III		
						{16}		
						I		

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3118214/table/T2/?report=objectonly>

Combination of microarray data from a total of 6 different experiments and 16 different time points {in parentheses} to analyse 4 different responses to mastitis infection: (I) overall response, (II) early stage response, (III) late stage response, and (IV) cattle-specific response.

themes, including the involvement of genes and pathways that were not identified in individual studies, using limited set of health data. These results pave the way for the development of novel therapeutics and vaccines for mastitis in ruminants.

Combination of microarray data from a total of 6 different experiments and 16 different time points {in parentheses} to analyse 4 different responses to mastitis infection: (I) overall response, (II) early stage response, (III) late stage response, and (IV) cattle-specific response.

Best health data for general health issues are likely those which can allow the best means of describing host genetic effects. However, the following dilemma will be soon faced, namely, the most precisely defined health phenotypes will allow a better detection power of genetic factors involved in the related health traits, but as shown previously, these narrowly defined health phenotypes are often the most difficult to measure and collect.

Collection of health data thus needs to be optimized, also by carefully assessing an optimum timing for the different measures. Advance knowledge in the kinetics of different process of the host responses (e.g. immune responses) and the individual variability of the host in mounting the responses will have to be acquired (e.g. in prior pilot studies) to make best use of health data.

Best health data may be also indirect trait data, for example in case of stress induced immune suppression like some subclinical mastitis cases, where collecting stress data might be as informative as direct pathological data.

Several research areas needing health data are emerging or finding new prospects with the availability of high throughput technologies. Developing novel vaccines is one of them: ideally, integrated approach to disease control aims at identifying genes responsible for the genetic variation, both of disease resistance and of response to vaccination. Such research is developed for viral diseases in cattle like bovine respiratory disease (Glass *et al.*, 2012) and can provide novel tools to select more resistant animals and more efficient vaccines. If early step of vaccine development may use controlled experiment data, further measures of vaccine effectiveness will necessarily need sufficient field health data.

Another emerging field is the analysis of the microbiota. There is increasing interest in studying not only the interactions between the host and the pathogen, but also to include the microbiota in this continuous interactive process. Measuring and analyzing microbiota data, as an integrated picture to monitor the health status of animals will become essential, especially for gut or respiratory health.

---

---

### What are the best health data?

---

---

---

---

### Prospects for using health data

---

---



---

---

## Conclusion

---

---

Research and development aiming at understanding health related mechanisms and implementing integrated strategies for improved health of livestock require large use of recorded health data, collaborative multidisciplinary programs and efficient technology transfer between research and industry.

If we look towards innovation, the key challenges for collecting and using health data will be likely among the following recommendations:

- Integrate health related traits in existing multi-trait selection programmes,
- Aim to use the same health traits for breeding and management,
- Develop cost effective tools to analyse field disease outbreaks and develop predictive diagnostics,
- Support initiatives integrating genomic approaches to vaccine development,
- Facilitate research including collaborations between animal health research institutions, commercial breeding and pharmaceutical companies,

with a common objective of joint acquisition of health data in order to generate new knowledge and innovation.

---

---

## List of references

---

---

Bai, Y., M. Sartor & J. Cavalcoli, 2012. Current status and future perspectives for sequencing livestock genomes. *J of Animal Science and Biotech* 3(1): 8. doi: 10.1186/2049-1891-3-8

Bishop, S. C., A.B. Doeschl-Wilson & J. A. Woolliams, 2012. Uses and Implications of Field Disease Data for Livestock Genomic and Genetics Studies. *Front Genet.* 3: 114. doi: 10.3389/fgene.2012.00114

Boddicker, N., E. H. Waide, R. R. R. Rowland, J. K. Lunney, D. J. Garrick, J. M. Reecy & J. C.M. Dekkers, 2012. Evidence for a major QTL associated with host response to Porcine Reproductive and Respiratory Syndrome Virus challenge. *J Anim Sci*, 90: 1733-1746.

Davies, G., S. Genini, S. C. Bishop & E. Giuffra, 2009. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. *Animal* 3, pp 415-436. doi:10.1017/S1751731108003522.

Genini, S., B. Badaoui, G. Sclep, S. C. Bishop, D. Waddington, M.H. Pinard van der Laan, C. Klopp, C. Cabau, H. M. Seyfert, W. Petzl, K.Jensen, E. J. Glass, A. de Greeff, H. E. Smith, M. A. Smits, I. Olsaker, G. M. Boman, G. Pisoni, B. Castiglioni, P. Cremonesi, M. Del Corvo, E. Foulon, G. Foucras, R. Rupp & E. Giuffra, 2011. Strengthening insights into host responses to mastitis infection in ruminants by combining heterogeneous microarray data sources. *BMC Genomics* 11;12(1):225. doi: 10.1186/1471-2164-12-225.

Glass, E.J., R. Baxter, R.J. Leach & O.C. Jann, 2012. Genes controlling vaccine responses and disease resistance to respiratory viral pathogens in cattle. *Vet Immunol Immunopathol.* Jul 15;148(1-2):90-9. doi: 10.1016/j.vetimm.2011.05.009.

Hamzic, E., B. Bed'Hom, H. Juin, R. Hawken, M. Abrahamsen, J.M. Elsen, B. Servin, M.H. Pinard-van der Laan & O. Demeure, 2013. Investigation of immune response to *Eimeria maxima* in broilers. 8th European Symposium on Poultry Genetics, 2013, Venezia, Italy.

Kirkpatrick, B. W., X. Shi, G.E. Shook, & M.T. Collins, 2011. Whole-Genome association analysis of susceptibility to paratuberculosis in Holstein cattle. *Animal Genetics* 42: 149-160. doi: 10.1111/j.1365-2052.2010.02097.x