European Journal of Epidemiology **16**: 601–606, 2000.

Herd immunity and herd effect: new insights and definitions

T. Jacob John[1] & Reuben Samuel[1,2]
[1]*Department of Clinical Virology, Christian Medical College Hospital;* [2]*Department of Community Health, Christian Medical College, Vellore, Tamil Nadu, India*

Accepted in revised form 21 April 2000

"Words are not only vehicles to convey ideas but also their drivers."

Anon.

Abstract. The term herd immunity has been used by various authors to conform to different definitions. Earlier this situation had been identified but not corrected. We propose that it should have precise meaning for which purpose a new definition is offered: "the proportion of subjects with immunity in a given population". This definition dissociates herd immunity from the indirect protection observed in the unimmunised segment of a population in which a large proportion is immunised, for which the term 'herd effect' is proposed. It is defined as: "the reduction of infection or disease in the unimmunised segment as a result of immunising a proportion of the population". Herd immunity can be measured by testing a sample of the population for the presence of the chosen immune parameter. Herd effect can be measured by quantifying the decline in incidence in the unimmunised segment of a population in which an immunisation programme is instituted. Herd immunity applies to immunisation or infection, human to human transmitted or otherwise. On the other hand, herd effect applies to immunisation or other health interventions which reduce the probability of transmission, confined to infections transmitted human to human, directly or via vector. The induced herd immunity of a given vaccine exhibits geographic variation as it depends upon coverage and efficacy of the vaccine, both of which can vary geographically. Herd effect is determined by herd immunity as well as the force of transmission of the corresponding infection. Clear understanding of these phenomena and their relationships will help improve the design of effective and efficient immunisation programmes aimed at control, elimination or eradication of vaccine preventable infectious diseases.

Introduction

The term 'herd immunity' is increasingly frequently seen in recent literature on the epidemiology of infectious diseases and on their prevention and control by immunisation. This was found by Medline search (1992–1998) using the key words 'herd immunity' and 'herd effect'. While a large number of papers were found with the key word 'herd immunity', none was found with the key word 'herd effect'. A review of several papers has shown that 'herd immunity' is used by different authors for different ideas, such as a concept, a phenomenon or a measurable parameter. In this paper we offer new insights and suggest a new definition of the term 'herd immunity' and dissociate it from the indirect effect of immunisation of part of a population on the incidence of infection or disease in the unimmunised remainder, for which the term 'herd effect' is proposed and defined.

Definitions currently available for herd immunity

Three definitions of herd immunity given in recent text or reference books are quoted below. The emphasis of phrases is ours.

(1) "Herd immunity. The *resistance* of a group to attack by a disease because of the immunity of a large proportion of the members and the consequent lessening of the likelihood of an affected individual coming into contact with a susceptible individual" [1].

(2) "Herd immunity. It is not necessary to immunise every person in order to stop transmission of an infectious agent through a population. For those organisms dependent upon person-to-person transmission, there may be a definable *prevalence of immunity* in the population above which it becomes difficult for the organism to circulate and reach new susceptibles. This *prevalence* is called herd immunity" [2].

(3) "Herd immunity. It is well known that *not everyone in a population needs to be immunised to eliminate disease* – often referred to as herd immunity. This is because successful immunisation reduces the number of susceptibles in the population and this effectively reduces the efficiency with which the microbe is transmitted from one person to the other. This has the same effect on the incidence of infection as a reduction in the number of individuals in a population. The microbe cannot sustain itself and disease disappears at some level of vaccine

coverage that is less than 100%. On the other hand, coverage below that needed to prevent disease may have little impact on the total number of susceptibles – as has been predicted using mathematical models and verified, in the case of rubella, by observation. The implementation of immunisation programmes needs to be accompanied by case surveillance, in conjunction with analysis of appropriate serological samples both before and after the introduction of the vaccine. If such data are not carefully scrutinised, the consequence may be dire. For example, an immunisation programme may reduce the number of cases but at the same time may increase the average age at which the infection occurs. If the severity of disease increases with age of acquisition, as in rubella (the risk of an infected foetus in an infected woman) and polio (the risk of paralytic disease), an immunisation programme may be less useful than none at all" [3].

The first definition considers herd immunity conceptually as the resistance to disease due to reduced risk of infection in a group of individuals as a result of a large proportion among them (but not all) being immune, not necessarily due only to immunisation [1]. The second definition clarifies that the term applies to the actual proportion of immunised individuals necessary to make it difficult for the organism to circulate and reach new susceptibles [2]. By this definition herd immunity is a threshold value of a measurable parameter (vaccination coverage) resulting in retardation of person-to-person transmission of an infection. Elimination of infection or disease is not required but only implied in this definition. How can we arrive at a definable level (proportion) of immunity when the end point is merely difficulty of transmission, which is not defined, hence not measurable? The third definition is closely similar to the first but it develops it further to a phenomenon providing the means to eliminate an infectious disease from a population when a proportion that is less than 100% is immunised [3]. Here the phenomenon has a measurable end point – that of the disappearance or elimination of disease in a group. In contrast to the second definition, here the actual input (immunisation coverage) necessary to achieve this measurable endpoint can be quantified. However it seems that the term herd immunity refers to the phenomenon of zero incidence in the unimmunised segment rather than the actual proportion immunised to achieve it [3]. The strict application of this definition requires the disappearance of disease due to immunisation coverage of less than 100% of susceptibles. This is a very rare event, exemplified by the global eradication of smallpox without immunising all susceptibles and the elimination of poliomyelitis due to wild polioviruses form North America when only some 65 to 70% immunisation coverage had been reached [4]. Interestingly, in the case of poliomyelitis, near 100% coverage of infants and additional annual repetitive vaccination of the same cohorts of susceptible children were necessary for 8–9 years to interrupt transmission in Brazil, showing that 'herd immunity' was not evident for the same vaccine against the same disease, but in a different region [5]. This clearly shows that herd immunity (by the third definition) is a function of not only immunisation but also the force of transmission of the pathogen. This definition is ambiguous about one aspect of the end point: is it zero disease or infection? The second definition clearly addresses infection while the first one focuses on infection and disease. Only the second definition clarifies that we are dealing with only person-to-person transmitted infectious diseases.

A new insight in the third definition is that in some cases it might be harmful to immunise a proportion and obtain only partial or incomplete 'herd immunity', since it might worsen the problem by delaying infection and not eliminating it, as in the case of maternal and foetal rubella [2]. This is an important point. A similar situation might arise in the cases of immunisation against hepatitis A and varicella also. Partial coverage of children may slow down virus circulation and delay infection in the unimmunised. Hepatitis A and varicella are more severe in adults than in children. Here a new phrase such as 'partial' or 'incomplete' herd immunity had to be introduced to overcome the problem caused by the very definition which demands disappearance of disease.

A review of several recent publications with the key word herd immunity showed that the term has most often been used to mean the concept of reduced transmission due to high immunisation level, in accordance with one or another of the three definitions given above. For example the 'concept of herd immunity' has been advocated as useful in designing immunisation programmes. [6, 7]. Herd immunity has been qualified as a 'key concept' in population based immunisation programmes [7]. The author of a landmark review of the history, theory and practical aspects of herd immunity chose not to prefer "any single definition of herd immunity, rather accepting the varied uses of the term by different authors" [8]. Recognising the consequent potential for confusion, the reviewer coined the phrase 'herd immunity threshold' to indicate the minimum prevalence of immune individuals necessary to interrupt transmission of infection [8]. Its major purpose was to distinguish between the desirable outcome of interruption of transmission from the potentially undesirable effect mentioned above. In general herd immunity is considered desirable by most authors, but there is incongruity in using the term 'immunity' to cover adverse outcomes also. In summary, the definitions are not clear, precise or complete; nor do they agree among themselves. A precise definition is necessary. We do not favour the status quo approach adopted by one reviewer [8].

The proposed new definition for herd immunity

The term herd immunity contains two words, herd (meaning a group or community) and immunity which has to be interpreted. Previously the term immunity did mean a state of protection but today it means a state in which the immune system of the body has reacted specifically to defined immunogen(s). It is an attribute of the individual, not a group. If so desired immunity can be tested for, as the presence of antibody, or as skin test or lymphocyte response to stimulation, against the chosen antigen. Putting the two terms together, it is proposed that herd immunity be now defined simply as the proportion of subjects with immunity in a given population. Under this definition herd immunity is quantifiable by testing a sample of the population for the presence of the selected immune parameter. It may be due to natural infection, or immunisation, or a combination of both. It is not dependent on the ease or difficulty of circulation of an infectious agent, nor its elimination. In certain cases, past infection (inducing long lasting immunity) and immunisation (inducing near 100% immunity which is long lasting) may be used as surrogates of immunity for quantifying herd immunity.

Indeed, some authors have actually used the term herd immunity to mean the proportion of subjects with immunity, in conformity with our new definition and clearly at variance with the definitions cited earlier [9–11]. They have obviously assumed that they were using the term correctly [9–11]. For example, in a survey of antibody prevalence by age to varicella-zoster virus, the prevalence was referred to as herd immunity [9]. In a study of measles outbreaks in two adjacent towns in Japan, the term 'herd immunity level' was used interchangeably with the sum of frequency of vaccination and of previous history of measles [10]. The authors presented it as the complement of 'susceptibility rate' [10]. In another study of antibody prevalence and hepatitis A outbreaks, the former was equated to herd immunity [11]. There are many more such publications in which the term herd immunity has been used in accordance with our new definition.

To recapture the spirit of the earlier but imprecise definitions, we can say that as herd immunity due to immunisation increases, at some point, which is short of 100% coverage, the circulation of the corresponding agent may cease. That point or value of herd immunity (called 'herd immunity threshold' by one author) cannot be the same for different diseases and for the same diseases it need not be the same in different communities [8]. An interesting question could be asked: can the consequence (short of elimination of infection) of increasing herd immunity by immunisation be measured? If it can be, then the fixation on the requirement of interruption of transmission can be eliminated. For this purpose a different term, namely 'herd effect' is introduced to denote the perturbation, if any, on the incidence of disease or infection in the unimmunised segment of a population, induced by the herd immunity of immunisation.

The proposed definition of herd effect

It is proposed that herd effect be defined as the alteration of the epidemiological frequency parameters (of infection or disease as the case may be) in the unimmunised segment of a population as a result of immunising a proportion of the population. The alteration is usually a decline of incidence of infection (hence lower incidence of disease). Therefore it may be simpler to define it as the reduction of infection or disease in the unimmunised segment as a result of immunising a proportion of the population. Lowered incidence of infection due to the herd effect of immunisation may, in some cases, be associated with increased incidence of disease consequent upon an upward shift in age of infection (example hepatitis A, maternal rubella syndrome in infants) or increased severity disease (such as adult varicella). However, in the age group immunised there will always be a reduction of incidence if there was herd effect. In most, if not all other cases, herd effect of high herd immunity induced by immunisation is beneficial in reducing the burden of disease; its extreme benefit is the interruption of transmission itself. In some earlier publications by one of us, the term herd effect had actually been used with this meaning [12, 13].

Earlier we had drawn attention to the contrast between North America where polioviruses were eliminated with routine immunisation coverage (with 3 doses of oral poliovaccine) of below 80% and South America where near 100% coverage with some 9–10 doses were required for the same effect [5]. According to the new definitions the herd immunity and the herd effect were different in these two regions. The reason for dissimilar herd immunity is the geographic difference in vaccine efficacy, necessitating a higher level of vaccine coverage in the region with lower vaccine efficacy to achieve a similar level of herd immunity. The large difference in the vaccine coverage and number of doses needed to eliminate polioviruses in South America in comparison to North America cannot be attributed to the lower herd immunity alone but to insufficient herd effect as well. The difference in herd effect for similar levels of herd immunity is due to difference in the force of transmission of wild polioviruses [12, 13]. In other words, when 100 children are given three doses, whereas in North America virtually 100% herd immunity is induced, in South America and in India only some 70% is induced, proving lower vaccine efficacy [12, 13]. Whereas the median age of poliomyelitis (in the pre-immunisation era) in North America was above 15 years, it was below 15 months in India, proving higher force of transmission [12, 13]. The confounding of the assumed direct relationship between vaccine coverage and incidence of infection by these two phenomena (dissimilar herd immunity for similar vaccine coverage and dissimilar herd effect for similar herd immunity) confused many world experts on polio, who persisted with the three-dose dogma (until about 1990), mainly because of the lack of clear definition of the term herd immunity and the lack of understanding of the relationships between herd immunity, herd effect and transmission of polioviruses. As an aside, the enhanced potency inactivated poliovaccine does not show any geographic variation of vaccine efficacy. We have observed much better herd effect of the latter vaccine; this is additional evidence for our argument presented above [12, 13], In short, when herd immunity is redefined to mean the immunity prevalence in a population, the cessation of transmission due to immunisation can be seen as the herd effect of high herd immunity due to high immunisation coverage with a vaccine having high vaccine efficacy.

Since herd effect is not necessarily break in transmission, but the reduction in transmission, a lower coverage level would have a lower herd effect and a very high coverage level could result in high herd effect leading to zero incidence even in the unimmunised. While herd immunity is applicable to any infection irrespective of its transmission pathways and to immunity induced naturally (by infection) or artificially (by immunisation), herd effect applies only when infected persons participate in the transmission of an agent and when immunisation induces at least some protection against infection (and not merely against disease). Thus, immunisation against tetanus or rabies (even if given routinely) will have no herd effect. As BCG inoculation seems to protect only against progressive primary tuberculosis and not against secondary type pulmonary tuberculosis, it also has no herd effect.

The measurement of herd effect

To be useful in understanding and quantifying the effects of immunisation it is not sufficient to differentiate herd effect from herd immunity, but we should be able to measure them. It was pointed out earlier that herd immunity is measurable by testing a sample of the population for the required parameter of immunity. Under certain circumstances, herd effect can also be measured. One of the best studies in which it was measured is that of immunisation of children with pertussis toxoid decreasing the spread of pertussis within the family [14]. The investigators found that vaccination of children with the toxoid was followed by significant reduction in incidence of pertussis among siblings and parents. They did not use the term herd effect, but called it 'indirect protection'. Interestingly they measured this effect as: (Indirect protection) = $(1 - R)$, where R was the ratio of incidence rates in contacts (parents and siblings) of recipients of toxoid (immunised) and of placebo (unimmunised) [14]. This parallels the method of measurements of vaccine efficacy, (VE) = $(1 - R)$, where R is the ratio of the incidence rates in immunised and unimmunised. Therefore, for ease of use, we may use the term 'VE equivalent' for the measure of herd effect, since the method of measurement and the way it is expressed as the percent reduction in incidence are both very much like VE itself.

It now becomes obvious that herd effect confounds the measurements of VE, since the incidence in the unimmunised is altered by the introduction of immunisation in a group. Thus, the measurement of VE subsumes herd effect when it is estimated in 'field' (community) immunisation programme settings. An interesting corollary of this relationship between herd effect and VE is that in a community in which immunisation has already been introduced the true VE cannot be measured by comparing the incidences in the vaccinated and the unvaccinated; the measured VE will be an underestimate. In such situations either historical data on incidence prior to introducing vaccination or incidence in an unvaccinated but similar community must be used to obtain the true measure of VE, corrected for herd effect. If VE is measured in a clinical study, in which a limited number of children living in a large community are immunised, implying thereby that herd effect is minimum or none, then we may get a different value. If the immune response of immunised children is measured in the clinical setting, then we get only a surrogate for VE without exhibiting protection, and not confounded by herd effect. For these reasons, the term vaccine efficacy should be qualified as 'immunogenic VE,' 'clinical VE' and 'field level VE' as has already been suggested [12].

Relevance of herd effect and herd immunity in immunisation programmes

From the definitions of and relationships between herd effect and herd immunity stated above it is clear that neither is relevant for the vaccination of an occasional individual, but both are important elements in large scale or community wide immunisation programmes. Programme managers are advised to factor in these issues in planning an immunisation programme, or when introducing a new vaccine in an existing programme, and also to monitor herd immunity and herd effect in the community as part of planned assessment/evaluation. The relationship between the two will depend on the type of immunity in consideration, its realised value at a point in time and the rate at which it is realised. Immunising agents which provide protection only against disease but not against infection may have no herd effect. Many vaccines protect well against disease but not so well against infection. Since infection without disease is silent, it would seem that the protection is against infection also. Partial protection against infection may result in lower extent of multiplication or shorter duration of potential transmissibility, thereby tending to reduce the risk of infection in others. The level of herd effect induced by a vaccine (providing at least some protection against infection) will depend on the level of herd immunity at a point in time, probably in a non-linear fashion and this would further be modified by the rate at which this level is achieved in relation to time. For example, we have shown that the herd effect achieved in 'pulse immunisation' is very much higher than that of routine immunisation programme, for the same level of herd immunity [12, 13, 15]. In our pioneering experiment, with 65% coverage with 3 doses of OPV given in pulse fashion, the transmission of poliovirus appeared to have ceased abruptly, in contrast to its continued circulation in spite of over 90% coverage in year round routine immunisation [15]. Even though this difference may be partly due to the short duration of gut immunity induced by OPV, the principle of excellent herd effect of pulse immunisation is still valid as has been shown with successful elimination of measles from a community by annual single day pulse immunisation of only children over one year [16].

Investigation, clear conceptualisation and application of these relationships will improve our ability to design effective and efficient immunisation programmes aimed at control, elimination or eradication of vaccine preventable infectious diseases in the transmission of which humans participate significantly.

Herd effect of interventions other than vaccination

It is worthwhile to realise that herd effect is the consequence of reducing transmission. Therefore other interventions which also reduce the transmission potential will have similar herd effect. To control lymphatic filariasis, mass treatment with Ivermectin or diethyl carbamazine is being used in many endemic areas. By reducing the parasitaemia in individuals and reducing the number of parasitaemic individuals, the drug causes a reduction in transmission to susceptibles even if they had not taken the drug. Mass application of the drug is akin to pulse application of a vaccine. If done well, filariasis could be eliminated by this approach since there is no extra-human source of infection. Even though 100% of population do not (and need not) receive treatment, the reduction of infection in the untreated segment is the herd effect of pulse therapy with the drug.

The public health approach to tuberculosis control is the early detection and antimicrobial treatment of pulmonary tuberculosis in the hope that further spread may be curtailed. This is another example of 'indirect protection' not due to immunity but due to decreased chance of spread, very much like in the case of reduced transmission due to immunisation. If increasing proportions of persons with pulmonary tuberculosis are diagnosed and treated early, the incidence of infection in the susceptible population should continue to decline; this may also be called herd effect. On the other hand, if multi-drug therapy of leprosy should have similar herd effect, we must be sure that transmission occurs from the infected persons directly (like tuberculosis) or via vector (like filariasis). The lack of conclusive evidence for such transmission is a lacuna in the strategy to control leprosy taking advantage of the assumed herd effect of multi-drug therapy. In the case of malaria, chemoprophylaxis in a large proportion of persons and/or the use of insecticide-impregnated bed nets may also have some herd effect. This is due to the fact that each prevented infection reduces further transmission and also because mosquitoes may not get access to infected persons sleeping in nets.

Thus herd effect may in general be defined as the change (reduction) in the incidence of infection or disease in the un-intervened segment of a susceptible population due to the intervention in the rest of the population, compared to the incidence in the absence of intervention in the entire population. However, we have focused on the herd effect of immunisation programmes in this paper and we recommend that the term be qualified by the nature of intervention when appropriate (as due to early diagnosis and treatment, mass treatment etc). In ordinary use, when not qualified, herd effect will be taken as due to herd immunity.

References

1. Saunders WB. Dorland's Illustrated Medical Dictionary. 28th Edn. Philadelphia: 1994; p 812.
2. Keusch GT, Bart KJ. Immunization and Vaccine Use. In: Fauci AS, Brawnwald E, Isselbacher KJ, et al. (eds), Harrison's Practice of Internal Medicine. 14th Edn, New York: McGraw Hill, 1998; pp 758–771.
3. Isaacs D, Moxon ER. Immunisation. In: Weatherall DJ, Ledingham JGG, Warrel DA (eds), Oxford Text Book of Medicine. 3rd Edn, Vol. 1. Oxford: Oxford University Press, 1996; p 315–321.
4. Nathanson N. Eradication of poliomyelitis in the United States. Rev Infect Dis 1982; 4: 940–945.
5. De Quadros CA, Andrus JK, Olive JM, Carrasco P. Strategies for poliomyelitis eradication in developing countries. Public Health Rev 1993/94; 21: 65–81.
6. Anderson RM. The concept of herd immunity and the design of community-based immunisation programmes. Vaccine 1992; 10: 928–935.
7. Anderson RM, Donnelly CA, Gupta S. Vaccine design, evaluation and community-based use for antigenically variable infectious agents. Lancet 1997; 350: 1466–1470.

8. Fine PEM. Herd Immunity: History, theory, practise. Epidemiol Rev 1993; 15: 265–302.
9. Ooi PL, Goh KT, Doraisingham S, Ling AE. Prevalence of varicella-zoster virus infection in Singapore. South East Asian J Trop Med Public Health 1992; 23: 22–25.
10. Ohga S, Okada K, Miyazaki C, Akazawa K, Ueda K. The measles outbreak in Chikuhau district, Fukuoka, Japan, 1990: Correlation between herd immunity level and outbreak size. Acta Paediatr Japan 1992; 4: 447–453.
11. Rooney PJ, Coyle PV. The role of herd immunity in an epidemic cycle of hepatitis A. J Infect 1992; 24: 327–331.
12. John TJ. Immunisation against polioviruses in developing countries. Rev Med Virology 1993; 3: 149–160.
13. John TJ. Can we eradicate poliomyelitis? In: Sachdev HPS, Choudhury P (eds), Frontiers in Paediatrics, New Delhi: Jaypee Brothers, 1996; p 76–90.
14. Trollfors B, Taranger J, Lagergard T, Sundh V, Bryla DA, Schneerson R, Robbins JB. Immunisation of children with pertussis toxoid decreases spread of pertussis within the family. Pediatr Infect Dis J 1998; 17: 196–199.
15. John TJ, Pandian R, Gadomski A, et al. Control of poliomyelitis by pulse immunisation in Vellore, India. Brit Med J 1983; 286: 31–32.
16. John TJ, Ray M, Steinhoff MC. Control of measles by annual pulse immunization. Amer J Dis Child 1984; 138: 299–300.

Address for correspondence: Prof T.J. John, Department of Clinical Virology, CMC Hospital, Vellore – 632 004, Tamil Nadu, India