Contact Tracing – Old Models and New Challenges

Johannes Müller\textsuperscript{1,2}, Mirjam Kretzschmar\textsuperscript{3}

\textsuperscript{1} Mathematical Institute, Technische University München, Boltzmannstr. 3, 85748 Garching, Germany
\textsuperscript{2} Institute for Computational Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany
\textsuperscript{3} Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584CX Utrecht, The Netherlands

Abstract

Contact tracing is an effective method to control emerging diseases. Since the 1980’s, mathematical modelers are developing a consistent theory for contact tracing, with the aim to find effective and efficient implementations of contact tracing, and to assess the effects of contact tracing on the spread of an infectious disease. Despite the progress made in the area, there remain important open questions. In addition, technological developments, especially in the field of molecular biology (genetic sequencing of pathogens) and modern communication (digital contact tracing), have posed new challenges for the modeling community. In the present paper, we discuss modeling approaches for contact tracing and identify some of the current challenges for the field.

1 Introduction

Emerging and re-emerging infectious diseases like SARS, Ebola, Lassa fever, Tuberculosis, and most recently SARS-CoV-2, require rapid responses and targeted control measures. It is best if an immunization of the population is possible - however, in case of emerging diseases, often the necessary vaccines are not available yet, or not available in sufficient quantities. An alternative approach is to stop infection chains by non-pharmaceutical control measures, such as reducing infectious contacts by social distancing, and testing and isolating infectious individuals. This can be done by screening if sufficient testing facilities are available, and in addition tracing and quarantining contact persons of infected index cases [58].

Mass screening as a stand-alone control measure is effective if prevalence is high, and cheap, rapid, and reliable diagnostic procedures are available. Otherwise, most tested persons are uninfected such that even a small probability for a false positive test result leads to a large number of false alarms, whereas only few infected persons can be identified.

Contact tracing (CT) is a more focused method: Once an infected individual is diagnosed and isolated, contact persons are identified who had potentially infectious interactions with that index case. In general, the prevalence within this group will be much higher than that in the general population, such that it may be effective to screen these persons, and to quarantine or isolate them. CT acts on several levels of the transmission process: Individual level: Infected persons are diagnosed early and are isolated or receive medical help. Population level: Transmission chains can be detected and stopped, which reduces the effective reproduction number. Medical/scientific level: It is possible to study infector–infectee pairs, and to learn about who infected whom in the outbreak. This gives information on risk factors, transmission modes, infectivity, and generation
intervals of the infectious disease. This information can be central in the design of further control measures.

The main challenge in mathematical modeling of CT is the individual-based character of the process. Information about the health status of single individuals and the time course of the contacts between those individuals is necessary. An appropriate formulation of the process is only possible at the microscopic level. However, the main interest is at the mesoscopic and macroscopic level: Not single individuals but the spread of the infection is in the focus of the interest. Model approaches need to span several scales, from the individual to the population. Similar to CT, also the transmission of infection is a process occurring between individuals. That process can be readily approximated by mean field models as the Kermack-McKendrick model. For the epidemic process, it is well established how to bridge the scales. CT, in contrast, follows the interaction given by transmission events, so can be viewed as a kind of superinfection: One could consider CT as an infection that follows the paths of the primary infection, and removes in that way infected individuals. This process superimposed on the transmission process creates a high degree of dependency between individuals. Therefore, lifting a model description for CT from the individual level to the population level involves more technical difficulties than for the case of a pure infection process.

Intuitively, the CT and transmission processes are racing each other. Starting at an index case, the infection spreads to contacts, while with CT we aim to catch up. This picture already indicates many of the properties that make a disease controllable by CT: (a) a sufficiently large fraction of cases develops symptoms and is tested and diagnosed; (b) contacts need to be well identifiable and traceable; (c) the disease spreads slowly enough to allow CT to catch up, even if identification, testing, and quarantining of contacts comes with a certain delay, and (d) diagnostic tests are able to readily identify symptomatic and asymptomatic persons.

Here, we review published literature on mathematical models for CT and their applications. We then identify and discuss open problems and current challenges for the theory of CT.

2 Modelling approaches

As a rigorous mathematical analysis of detailed models for CT is rather challenging, a multitude of simplifying approaches have been established. Among the first papers is a study by Hethcote and Yorke concerning interventions for gonorrhoea. They include CT in their model and suggest that the effect of CT is a reduction of the transmissibility of the infection. In the same spirit, proposes that CT is based on the identification of infected persons by their infected contacts. In that, the “superinfection” is taken verbally, and infecteds are removed with a term that resembles the incidence term (product of discovered infecteds and prevalence). In those models, the term representing CT is not based on first principles, but rather defined ad hoc.

For models, that are based on first principles, three main directions can be identified:

1. Simulation models that directly simulate individuals and large populations;

2. Pair approximation models that are extensions of mean field models incorporating information about correlations of pairs of individuals;

3. Stochastic and deterministic models that are based on a rigorous analysis of a simple branching process modeling CT.
We discuss these in the following sections, and also the phenomenological approaches, which are relevant for practical applications.

2.1 Individual based simulation models

Individual (or agent) based models (IBMs) are perhaps the first choice to formulate a process as complex as CT. Individual-based models describe the fate of every single individual in a population and their interactions. IBMs explicitly incorporate a contact graph, which can be as simple as a complete graph where every individual has contact with every other individual (fully connected population), a random graph as described by the configuration model, or a small world graph that reflects local and long distance contacts. Conceptual, parsimonious models are used to address the influence of the contact graph structure on transmission dynamics.

The most detailed IBMs describe graphs that aim to represent existing societies with cities, workplaces, and schools. It depends on the aim of the model how detailed the contact graph, the state of an individual, and its behavior is formulated. As such a model is algorithmic, there is almost no limit for the degree of detail that can be included. However, very detailed models are often faced with the problem of lack of data for an appropriate parametrization.

In any case, the advantages of individual-based models for CT are three-fold: First, they are simple to formulate and to communicate; second, they allow to directly represent CT in a realistic way and in that foster and guide the development of more abstract, analytical models; and third, they provide sufficient detail, such that results can be used in public health decision making.

IBMs are frequently used to investigate specific infectious diseases, and to extract relevant information about the effectiveness of intervention strategies. Contact tracing is commonly performed to improve case finding for sexually transmitted diseases as gonorrhoea and chlamydia; for these diseases, contacts can be clearly defined - recent sex partners - and a large fraction of infections are asymptomatic, such that contact tracing greatly enhances the possibility of finding and treating infected persons. Also, the time scale of transmission of these infections is sufficiently slow to enable finding and treating contacts on a faster time scale than the generation interval. Finally, as sexual partnerships are often long lasting, treating contacts may prevent re-infections of the index case. Several IBM models investigate Tuberculosis, also in connection with bioterrorism, measles, and Ebola. Also, the effect of CT on control of SARS has been simulated by various IBMs.

2.2 Pair approximation models

Stochastic processes are hard to analyze, while many tools for deterministic models are available. Mean field equations are a well established, heuristic approach to reformulate individual-based models as described above in terms of ordinary differential equations (ODE’s). In the limit, if the population size tends to infinity, the approximation becomes exact in case of a homogeneous population; that precisely is the way to derive a deterministic model as the Kermack-McKendrick (or SIR) model from a stochastic process for an infection. Instead of counting the number of individuals of different types, the expected relative frequencies are addressed by the ODE model.
individuals. In that, the expectations exactly satisfy an ODE. Other transitions, as transmission of infection, are based on interaction of individuals. At that point, only an approximation is possible that neglects all correlations in the contact graph. Formally, the expectation of the product of random variables is replaced by the product of expectations [85, chapter 3.1.10]. Under appropriate conditions, particularly if the contact graph is a complete graph (contacts happen between any of the individuals), this procedure becomes rigorous if the population size becomes large or infinite. The success of ODE models to describe the time course of real-world epidemics is an \textit{a posteriori} justification for simple deterministic models. In many cases, they allow for a deeper understanding of the underlying mechanism, and for powerful predictions [51].

All information about correlations are lost in the transition from an IBM to a mean field model. In that, a mean field model is insufficient to appropriately describe the infectious process on an inhomogeneous contact graph, particularly if the contact graph is strongly locally clustered: Is this the case, the neighbor of an infected individual often already is infected, and the spread of infection slows down. In the 1990’s, mainly driven by Japanese [99] and British [50] groups, an improved mean field approximation was developed, the pair approximation. Here, not only expectations but also the correlations are formulated in an ODE model. That is, not only the expected number of individuals in state \( S \), \( I \), or \( R \), say, but also the expected number of edges connecting e.g. \( I \) with \( I \) or \( S \) with \( I \) individuals are followed in the system of ODE’s. As the model incorporates information about correlation, it can – up to a certain degree – mimic the slow down of the spread caused by spatial correlations. The disadvantage of this approach is the exploding number of equations needed, as not only one equation per state is required, but also one for each pairwise combinations of two states. Therefore, most authors do not perform an analysis based on the theory of dynamical systems, but the resulting ODEs are solved numerically. Furthermore, it turns out that – for strongly localizing graphs – the conclusions are rather of qualitative than quantitative nature. If the contact graph is well mixing (high degree of the nodes, and no clustering), the pair approximation becomes better (which is also the case for the mean field approximation). Recent developments allow even an exact analysis by the “message passing method”, a subtle further development of the pair approximation [48, 102, 110] for trees. Concerning the modeling of CT, pair approximation keeps a central piece of information that is dismissed by the mean field approximation: We know how likely it is that a neighbor of an infected individual is infected. In that, it is possible to remove infected neighbors of an index case [25, 14]. This idea has been discussed in a series of papers [108, 26, 52, 40, 41]. The comparison with Monte Carlo simulations of the stochastic process indicates that the results are valid especially for large, homogeneous graphs. Most of these papers investigate the effectiveness of CT in different contact graph structures. In [40] it is found that CT is more effective in clustered than in homogeneous populations. In recursive and one-step tracing is compared, and “targeted CT”, that is, CT focusing on a risk group, is analyzed. Recursive and targeted CT were found to be particularly effective. The models for CT based on pair approximation are rather conceptual models that allow addressing fundamental questions than models used to quantitatively predict the effect of CT on the spread of real-world infections. One of the few exception is [18], where pair approximation is applied to predict the impact of CT on chlamydia prevalence. A simpler approach, in which only pairs are taken into account, which may form and dissolve, is used in [36] to investigate the impact of screening and CT, again addressing the prevalence of chlamydia.

2.3 Models based on branching processes

At the onset of an outbreak spreading in a homogeneous population, it is possible to replace the epidemic process by a birth-death process of independent individuals, where “birth” means
a new infection and “death” recovery. Particularly the interaction of two infectious individuals is unlikely and negligible in a large, homogeneous population \([6]\). The statistics of this process, e.g. the size and structure of connected components is known \([86]\). On top of this linear birth-death process, a process of CT can be formulated. It is possible to rigorously analyze this stochastic process \([84,7]\): CT mainly affects the removal rate of infected individuals. That is, in order to address the stochastic process, the probability to be infectious at a given time after infection is determined. This probability is the central function that allows to readily determine the effective reproduction number, or the doubling time of an infection. Various aspects of CT can be investigated in this context, as the recovery and infectivity depending on the time since infection \([84,7]\), estimation of the tracing probability from data \([82,12]\), or the effect of a tracing delay \([8]\). Strictly spoken, the analysis and the results are valid only for the onset (or during the decay before extinction) of the outbreak, if direct contacts between infecteds are unlikely to happen. However, using heuristic arguments, the removal rate can be approximated also in the case of high prevalence, and a modified mean field equation has been proposed \([84]\).

Similar to pair approximation, the approach is not suited for a complex contact graph structure with small homogeneous clusters that only weakly interact. It is interesting that the central idea for the analysis of CT on the one hand, and the message passing methods used in recent versions of pair approximation on the other hand, bear a remarkable similarity. In \([89]\), the branching process analysis is generalized from homogeneous populations to populations with a prescribed contact graph, and ideas are discussed how to merge the branching process analysis and the pair approximation. Also \([59]\) investigates CT on a random graph, this time via an approximation using generating functions for he degree distribution of neighbors of randomly chose infected individuals. They find, that the degree of an individual detected by backward tracing roughly behaves as the expectation of the squared degree, indicating the high efficiency for CT to detect the persons that are best connected.

Brown et al. \([14]\) developed a sophisticated ODE-approximation of the branching-process structure and applied that to Ebola. Becker et al. \([9]\) also proposed a simplified version of the model and investigated the SARS epidemic in a deterministic model with household-structure. Kretzschmar et al. \([62]\) used a branching process to model ring vaccination for smallpox. This model was recently developed further to investigate the effectiveness of CT for SARS-CoV-2 \([65,64]\).

Based on a branching-process formulation for CT, Tanaka \([104]\) analysed data for SARS-CoV-2 to estimate the fraction of asymptomatic cases.

### 2.4 Phenomenological approaches

Phenomenological approaches are not rooted in an analysis of stochastic processes that obviously model CT in an adequate way. In that, it is difficult to understand if the terms chosen to model CT are adequate. The advantage of these models is their simplicity – they are ODE models with rather simple structure (concerning CT) and can be readily analyzed or simulated. In that, these models are suited to address real world epidemics. The difficult part is the interpretation of the results (again concerning CT), as the clear connection with first principles is not obvious. Often, the tracing part is either formulated as a linear removal term, or as a mass action term.

Many of these models are applied to the HIV infection, where mostly a mass action term is used to represent CT \([21,45,45,20,57,1]\), but sometimes also simply a linear term \([43]\). In \([20,42]\) several ways to model CT – linear, mass action, and a saturation function – are compared with data. Interestingly enough, results in \([42]\) indicate that a mass action term for CT is inferior to a linear term or a saturation function. In a similar spirit, Clarke et al. \([19]\) adjust a power low term for CT based on simulations from an IBM. Also models for Chlamydia \([35]\), Tuberculosis \([8]\), Smallpox \([17]\), Ebola \([11]\), and models for SARS-CoV-2 \([34,73]\) are based on
Fraser [30] proposed an idea with a slightly more profound connection to CT. The model is based on age-since-infection. In that, it is similar to the branching-process analysis, but CT is formulated as a linear effect. In that article, characteristics in the timing between onset of symptoms and infectivity are identified that make an infection controllable by CT. Chen et al. [16] took up that approach to analyze a model for the SARS infection and Ferretti et al. [29] applied the model to the SARS-CoV-2 epidemic in Italy.

3 Challenges

The efforts of the last 40 years to develop a toolbox for CT have delivered a considerable amount of modelling approaches and results. There is a general agreement about a fundamental model structure, that can be readily realized in IBMs, and there are various ways to analyze the stochastic process either approximately or rigorously – at least, if we stick to simple models. Many simple ad-hoc models are published. Nevertheless, some of the central questions are still open.

3.1 Classical questions

Some questions and problems have been debated for quite a while. We pick a few of these “classical questions” which we consider as interesting and/or of practical need, and discuss them.

Modeling CT. Obviously, a multitude of models are in use. Some models, as the stochastic IBMs, directly simulate CT. In an IBM, it is easy to incorporate CT appropriately. Stochastic simulation models have the advantage that their outcomes are easy to interpret, but the disadvantage that they cannot be analyzed analytically. To obtain insight into the parameter dependence of CT and to develop general rules for CT is not straightforward or may be impossible if we exclusively rely on IBMs.

Other approaches, as the pair approximation models or models based on branching processes, use first principles to develop the model structure. Therefore, also these models are able to reflect CT in an appropriate manner. Even if approximations are used to derive simplified mathematical structures, it is straightforward to check their accuracy (compare the analytical results to simulations of the original models). In that, these models are appealing. However, in some cases, the derivation is rather technical and not straightforward to communicate. This class of models are well suited for theoretical considerations. For practical purposes, simpler model structures seem to be desirable.

At that point, the phenomenological approach comes in. Here, the models are mostly compartmental (ODE) models, as often used in ecology and epidemiology. These models can be readily analyzed and simulated. In that, practical applications of such simple models are easily possible, which is the strength of the phenomenological approach. The drawback is the fact that they are not rooted in first principles. It is hard to assess whether the model structure appropriately reflects reality. E.g., if CT is formulated as a linear term, this formulation is contradicting the observation that CT is based on correlations and dependencies between index cases and their infectees/infector (as a rule, dependencies are expressed by nonlinear terms). Another aspect is the parametrization of the models: Some central parameters can be obtained by observations on the micro scale. E.g., the probability to detect a contact can be estimated using the number of detected cases per index case [82, 12]. As phenomenological approaches dismiss the micro-scale and directly jump to the macro-scale, information that is available on the micro-scale is almost impossible to incorporate.
Easy to use models that are well accepted and approved by the modeling community are necessary but not yet in sight.

**Backward/Forward tracing.** Even in one of the very first papers about CT, the seminal work by Hethcote and Yorke [39], the distinction between backward and forward tracing is mentioned. Backward tracing means that the infector is detected by an infectee who becomes an index case, while in forward tracing the infector is the index case, while the infectee is detected. Since this first article, the relative importance of backward- and forward tracing is under discussion. An individual only has one infector, but in general several infectees. This fact might indicate that forward tracing is more important. On the other hand, if we randomly select an individual in a natural contact graph, the neighbor of this individual will on average have more contacts than an average individual. This finding, also called the “friendship paradox” in the context of social networks, has been proposed to be used in an early warning system for influenza [17]: Students were asked to name two friends. The authors of the study monitored the occurrence of influenza among those named persons. It turned out that the incidence of influenza increased two weeks earlier than it did among the average students. This observation indicates that backward tracing also is of importance, as most likely we will find persons who have many contacts and already had the chance to infect many of them. This is especially important for CT in sexually transmitted infections, which are often circulating in highly connected core groups.

The question of whether there are super-spreaders also falls within this context [71, 32, 77]. The frequency and the importance of super-spreaders for the dynamics of infection, and also for the impact of CT, has not yet been conclusively clarified; [58, 59] argue verbally that backward tracing is readily able to identify super-spreaders, s.t. a combination of backward- and forward tracing is efficient, also in the presence of super-spreader events. In [59], quantitative comparisons (based on analytic results for a branching-process model for CT on random trees) indicates that the effect of CT decreases with the variance of the degree distribution. This, in turn, is a hint that CT performs better without super-spreader events. Also simulation studies [56] point in that direction (at least if we consider the expected final size). The timing (latency and incubation period) of the infection itself will also have some influence. There is a need to investigate the effect of super-spreaders on CT more in depth.

**Endemic equilibrium.** CT is also performed if the infectious process is in its endemic equilibrium. In this case, in average each infected individual only has one infector and one infectee ($R_{eff} = 1$). Why does CT pay in that situation? For sexually transmitted diseases (STD), it is clear that the partner is of high risk, and a couple should rather be considered as a single entity. Partner notification in faithful pairs is a simple case of CT. It is more interesting to note that for some STDs asymptomatic persons can be infectious for a long time (several months). In that, particularly the tracing of asymptomatic persons may be decisive. CT is a method to find these highly infectious persons with possibly many contacts, who are difficult to localize otherwise.

**Tracing probability.** In order to monitor the effect of CT, it is desirable to estimate the tracing probability, that is, the fraction of identifyable contacts among all infectious contacts. While data for the number of detected cases are available, the number of missed cases is usually unknown. In [82,12], some statistical methods are developed to estimate the tracing probability. A related problem is the estimation of the abundance of asymptomatic cases from tracing data. [23] proposes a method based on household models for Yaws, a disabling bacterial infection, while Tanaka [104] aims to estimate the percentage of asymptomatic cases for SARS-CoV-2 using the branching process approach for CT. However, these questions are rather neglected by the recent literature and deserve a deeper investigation.
3.2 CT and genetic sequence data

In recent years, genetic sequencing of pathogen DNA became rather cheap, and genetic sequence data are readily available. Methods of population genetics are used to, e.g., estimate the prevalence of an infection [31]. Genetic sequence data is also used to identify clusters of infections, and even to identify and refine transmission trees (infector/infectee relations) within a cluster [97, 92, 22, 91]. However, the combination of data from CT and from sequencing techniques is hardly exploited by now. It is natural to ask what epidemiology could gain from that combination.

Clearly, from samples of infector-infectee pairs, the mutation rate of the pathogen can be estimated. Moreover, as an infection event forms a bottleneck for the population of the pathogen, the time since infection can be estimated, and in that, it is possible to narrow down the time of the infectious contact. A comparison of different infector/infectee pairs might help to sharpen the estimations for the prevalence.

The usage of genetic sequence data together with methods from population genetics is rather young, and many powerful methods – as the SMC [75] – are rather recent developments. We expect that useful tools become available in the near future.

3.3 Digital CT

The idea to use data from mobile phones to trace contacts is rather recent. First practical attempts to use digital sensors (RFID chips) to observe contacts and to investigate an empirically validated contact network in relatively small communities (school, hospital, conference) reach back to 2010 [95, 46, 103]. Soon it became clear that risk evaluation based on mobile phone data is possible and useful [68, 74]. Perhaps the first paper that considered digital CT (DCT) was an IBM-based simulation study by Farrahi et al. in 2014 [28]. Several manuscripts in the context of SARS-CoV-2 also address DCT [13, 15, 73, 29, 38, 64].

The ability to localize persons using data generated by smartphones already detects traffic jams [90], and is a source of major concern as these data could also be misused [105]. DCT could help to improve some crucial shortcomings of classical CT. The advantages of DCT are the rapid identification of infectious contacts, and the possibility to also rapidly inform contactees. In comparison with conventional CT, DCT has the potential to strongly reduce the tracing delay, and to increase the tracing probability. In that, infections may become controllable that were not controllable before. However, these potential benefits come with a number of technical, social, medical, and practical challenges. Contacts identified by mobile devices need to correlate with infectious contacts. A rapid and cheap test for the infection is necessary, as the number of persons to test will be much higher than in conventional CT. Aspects of privacy and data protection are crucial elements of DCT. Apps need to be accepted in the society. The concept of DCT, its strengths and its weaknesses have to be communicated clearly, in a way that citizens understand and accept DCT. The system can only work if a large part of the population participates in DCT.

In the present note, we rather focus on new modeling challenges posed by DCT, and less on the social and technical aspects, though all of these points are intertwined.

**Correlation between individuals.** Often, the mathematical analysis of CT models focus on the onset of the outbreak. In that, correlations (particularly between infected individuals) that do not come from an infector/infectee relation can be neglected. At present, conventional CT can afford to trace contacts only if rather few index cases are diagnosed. Once the number of diagnosed infected persons exceeds several hundred in a community, the effort to identify manually many contacts per person is not feasible anymore.
DCTS promises to overcome some of the logistic problems of CT and to allow for the identification of a huge number of contacts, even if a high number of infected individuals are present. The number of reported contacts per day and person in Europe is in the magnitude of $10^4$, with a large standard deviation (which is in the same range as the average number of contacts). If we have a tracing window of one week, we easily estimate 70 contact persons per index case (with a high variance). Depending on the nature of the infection, we can (and should!) immediately inform not only the direct contactees but also the contactees of the contactees (second level tracing). Note that second level tracing is dissimilar to the classical two-step tracing: In two-step tracing a contactee is first tested, and if the test is positive, he/she becomes an index case and the snowballing continues, while in second level tracing the contactees of contactees are immediately informed about their possible exposure, without waiting for a test result of contact persons. In that, we easily arrive at 500 direct and indirect contactees per index case. That is, the fraction of the population affected by DCT is around 500 times larger than the number of index cases, however, the number of actually exposed persons might be small. From a practical aspect, it is of utmost importance to identify features of an infection that imply the necessity of first- second- or even higher-level tracing.

The difference to the conventional procedure is the possibility for rapid and immediate information of direct and indirect contactees, without waiting for medical tests and diagnoses. Even in a moderate outbreak, it is likely that the groups of contactees related to different index cases will overlap. These effects tremendously complicate the mathematical analysis of the models. A clear challenge is the identification of techniques that allow for the mathematical analysis of this situation.

**High number of contactees / practical protocol.** As discussed above, DCT can identify a high number of contactees and inform them about direct or indirect infectious contacts. From practical considerations, a reduction of that number is desirable. Technical devices may give a score to each contact, and estimate the probability for infection. That might help, but it also might be the case that scores are not reliable. Most likely, we are faced with a serious practical problem: If all casual contacts are reported, many persons might receive a warning rather often. In that case, fatigue sets in and warnings are not taken seriously anymore. The DCT becomes useless. A possibility to escape this risk is the choice of a smaller subgroup from the set of contactees. One possibility could be simply a random sub-set. More effective is a classification of contactees in different risk groups, e.g. according to the number of recent contacts. Particularly asymptomatic persons that spread the infection for a longer time will appear among repeatedly identified contactees. Modeling approaches need to clarify the situation and to estimate the number of warnings a person receives under realistic conditions. In consequence, different strategies of reducing the number of contactees to be informed have to be defined. Models allow to estimate the impact of these strategies, and to filter out reasonable policies for DCT. Many parameters will influence the outcome, as the availability of tests, the social/contact structure, and the abundance of a tracing app on mobile phones.

**DCTS induce an inhomogeneity in the population.** The population is divided in a subpopulation with, and a subpopulation without a tracking device. In that, contacts within the DCTS-subpopulation are readily identified, while contacts between the two subpopulations or among the non-DCTS-subpopulation can only be identified by conventional CT. That situation is completely different from an imperfect test, which leads to a certain probability of undetected infectious contacts. It is likely that the acceptance of a tracing device correlates with social factors as education, political, or religious orientation. In that, connected subgroups appear that escape the detection by DCTS. These subgroups might form a reservoir for the infection, from which
importation to the general population may occur. The fraction of the population equipped with DCTS-equipment alone is not decisive, but also the distribution in a heterogeneous population. However, even for a homogeneous population, the division into two subgroups might lead to unforeseen effects. This aspect is a new one that deserves deeper investigation.

4 Conclusion

The joint effort of modeling and theoretical investigation of CT has borne many fruits, as the estimation of the impact of CT on the spread of infections, or methods to assess the influence of the tracing delay. Nevertheless, questions of practical importance still deserve our attention. New challenges, e.g. due to developments in genetic sequencing of pathogens and DCTS, have arisen. Simulation models promise first answers on a rapid time scale, but it will take some time until these new aspects are fully understood analytically, and this understanding is translated into practice.

References

[1] M. Agarwal and A. S. Bhadauria. Modeling H1N1 flu epidemic with contact tracing and quarantine. Internat. J. Biomath., 05(05):1250038, 2012.

[2] C. L. Althaus, J. C. Heijne, S. A. Herzog, A. Roellin, and N. Low. Individual and population level effects of partner notification for chlamydia trachomatis. PLoS One, 7(12), 2012.

[3] J. P. Aparicio and J. C. Hernández. Preventive treatment of tuberculosis through contact tracing. Cont. Math., 410:17–29, 2006.

[4] B. Armbruster and M. L. Brandeau. Contact tracing to control infectious disease: when enough is enough. Health Care Manag. Sci., 10(4):341–355, 2007.

[5] B. Armbruster and M. L. Brandeau. Optimal mix of screening and contact tracing for endemic diseases. Math. Biosci., 209(2):386–402, 2007.

[6] F. Ball and P. Donnelly. Strong approximations for epidemic models. Stochastic Processes Appl., 55:1–21, 1995.

[7] F. G. Ball, E. S. Knock, and P. D. O’Neill. Threshold behaviour of emerging epidemics featuring contact tracing. Adv. Appl. Probab., 43:1048–1065, 2011.

[8] F. G. Ball, E. S. Knock, and P. D. O’Neill. Stochastic epidemic models featuring contact tracing with delays. Math. Biosci., 266:23–35, 2015.

[9] N. G. Becker, K. Glass, Z. Li, and G. K. Aldis. Controlling emerging infectious diseases like SARS. Math. Biosci., 193(2):205–221, 2005.

[10] M. Begun, A. T. Newall, G. B. Marks, and J. G. Wood. Contact tracing of Tuberculosis: A systematic review of transmission modelling studies. PLoS ONE, 8(9):e72470, 2013.

[11] T. Berge, A. J. O. Tassé, H. M. Tenkam, and J. Lubuma. Mathematical modeling of contact tracing as a control strategy of Ebola virus disease. Internat. J. Biomath., 11(07):1850093, 2018.
[12] M. G. B. Blum and V. C. Tran. HIV with contact tracing: a case study in approximate bayesian computation. *Biostatistics*, 11(4):644–660, 2010.

[13] W. J. Bradshaw, E. C. Alley, J. H. Huggins, A. L. Lloyd, and K. M. Esvelt. Bidirectional contact tracing is required for reliable COVID-19 control. May 2020.

[14] C. Browne, H. Gulbudak, and G. Webb. Modeling contact tracing in outbreaks with application to ebola. *J. Theor. Biol.*, 384:33–49, 2015.

[15] V. B. Bulchandani, S. Shivam, S. Moudgalya, and S. L. Sondhi. Digital herd immunity and COVID-19. Apr 2020.

[16] S.-C. Chen, C.-F. Chang, and C.-M. Liao. Predictive models of control strategies involved in containing indoor airborne infections. *Indoor Air*, 16(6):469–481, 2006.

[17] N. A. Christakis and J. H. Fowler. Social network sensors for early detection of contagious outbreaks. *PLoS ONE*, 5(9):e12948, 2010.

[18] J. Clarke, K. A. J. White, and K. Turner. Exploring short-term responses to changes in the control strategy for Chlamydia trachomatis. *Computational and Mathematical Methods in Medicine*, 2012:1–10, 2012.

[19] J. Clarke, K. A. J. White, and K. Turner. Approximating optimal controls for networks when there are combinations of population-level and targeted measures available: Chlamydia infection as a case-study. *Bulletin of Mathematical Biology*, 75(10):1747–1777, Jun 2013.

[20] S. Clémenton, V. C. Tran, and H. de Arazoza. A stochastic SIR model with contact-tracing: large population limits and statistical inference. *Journal of Biological Dynamics*, 2(4):392–414, Oct 2008.

[21] H. de Arazoza and R. Lounes. A non-linear model for a sexually transmitted disease with contact tracing. *IMA J Math Appl Med Biol*, 19:221–234, 2002.

[22] A. M. Dennis, D. K. Pasquale, R. Billock, S. Beagle, V. Mobley, A. Cope, J. Kuruc, J. Sebastian, C. Walworth, and P. A. Leone. Integration of contact tracing and phylogenetics in an investigation of acute HIV infection. *Sexually Transmitted Diseases*, 45(4):222–228, Apr 2018.

[23] L. Dyson, M. Marks, O. M. Crook, O. Sokana, A. W. Solomon, A. Bishop, D. C. W. Mabey, and T. D. Hollingsworth. Targeted treatment of yaws with household contact tracing: How much do we miss? *Am. J. Epidemiol.*, 187(4):837–844, 2017.

[24] K. Eames. Contact tracing strategies in heterogeneous populations. *Epidemiol. Infect.*, 135:443–454, 2007.

[25] K. T. Eames and M. J. Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *PNAS*, 99:13330 – 13335, 2002.

[26] K. T. Eames and M. J. Keeling. Contact tracing and disease control. *Proc. R. Soc. B*, 270:2565–2571, 2003.

[27] M. Eichner. Case isolation and contact tracing can prevent the spread of smallpox. *American Journal of Epidemiology*, 158:118–128, 2003.
[28] K. Farrahi, R. Emonet, and M. Cebrian. Epidemic contact tracing via communication traces. *PLoS ONE*, 9(5):e95133, May 2014.

[29] L. Ferretti, C. Wyman, M. Kendall, L. Zhao, A. Nurtay, L. Abel-Dörner, M. Parker, D. Bousall, and C. Fraser. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, page eabb6936, 2020.

[30] C. Fraser, S. Riley, R. Anderson, and N. Ferguson. Factors that make an infectious disease outbreak controllable. *PNAS*, 101:6146–6151, 2004.

[31] S. D. W. Frost and E. M. Volz. Modelling tree shape and structure in viral phylogeny. *Phil. Transact. R. Soc. B*, 368(1614):20120208, 2013.

[32] A. P. Galvani and R. M. May. Dimensions of superspreading. *Nature*, 438(7066):293–295, 2005.

[33] A. C. Ghani, J. Swinton, and G. P. Garnett. The role of sexual partnership networks in the epidemiology of gonorrhea. *Sex Transm Dis*, 24:45–56, 1997.

[34] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. D. Filippo, A. D. Matteo, and M. Colaneri. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med.*, 2020.

[35] C. Heffernan and J. A. Dunningham. Simplifying mathematical modelling to test intervention strategies for chlamydia. *Journal of Public Health and Epidemiology*, 1(1):022–030, 2009.

[36] J. C. Heijne, C. L. Althaus, S. A. Herzog, M. Kretzschmar, and N. Low. The role of reinfection and partner notification in the efficacy of chlamydia screening programs. *Journal of infectious diseases*, 203(3):372–377, 2011.

[37] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, S. Funk, R. M. Eggo, F. Sun, S. Flasche, B. J. Quilty, N. Davies, Y. Liu, S. Clifford, P. Klepac, M. Jit, C. Diamond, H. Gibbs, and K. van Zandvoort. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health*, 8(4):e488–e496, 2020.

[38] E. Hernandez-Orallo, P. Manzoni, C. T. Calafate, and J.-C. Cano. Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: the case of COVID-19. *IEEE Access*, pages 1–1, 2020.

[39] H. W. Hethcote and J. A. Yorke. Gonorrhea transmission dynamics and control. *Lecture Notes in Biomathematics*, 56, 1984.

[40] T. House and M. J. Keeling. The impact of contact tracing in clustered populations. *PLoS Comput Biol*, 6:e1000721, 2010.

[41] T. House and M. J. Keeling. Insights from unifying modern approximations to infections on networks. *J. R. Soc. Interface*, 8:67–73, 2011.

[42] Y.-H. Hsieh, H. de Arazoza, R. Lounes, and J. Joanes. A class of methods for HIV contact tracing in Cuba: Implications for intervention and treatment. In *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*, pages 77–92. World Scientific, 2005.
13

[43] Y.-H. Hsieh, Y.-S. Wang, H. de Arazoza, and R. Lounes. Modeling secondary level of HIV contact tracing: its impact on HIV intervention in cuba. *BMC Infect. Dis.*, 10(1), 2010.

[44] R. Huerta and L. S. Tsimring. Contact tracing and epidemics control in social networks. *Phys. Rev. E*, 66(5), 2002.

[45] J. M. Hyman, J. Li, and E. A. Stanley. Modeling the impact of random screening and contact tracing in reducing the spread of hiv. *Math Biosci.*, 181(1):17–54, 2003.

[46] L. Isella, M. Romano, A. Barrat, C. Cattuto, V. Colizza, W. V. den Broeck, F. Gesualdo, E. Pandolfi, L. Ravà, C. Rizzo, and A. E. Tozzi. Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. *PLoS ONE*, 6(2):e17144, feb 2011.

[47] E. H. Kaplan, D. L. Craft, and L. M. Wein. Emergency response to a smallpox attack: The case for mass vaccination. *PNAS*, 99(16):10935–10940, jul 2002.

[48] B. Karrer and M. E. Newman. Message passing approach for general epidemic models. *Phys. Rev. E*, 82:016101, 2010.

[49] P. Kassie, J. R. Andrews, W. D. Kelton, and D. W. Dowdy. Timing of tuberculosis transmission and the impact of household contact tracing, an agent-based simulation model. *Am. J. Respir. Crit. Care Med.*, 189(7):845–852, 2014.

[50] M. Keeling, D. Rand, and A. Morris. Correlation models for childhood epidemics. *Proc. R. Soc. London, Ser. B*, 264:1149–1156, 1997.

[51] M. J. Keeling. Correlation equations for endemic diseases: externally imposed and internally generated heterogeneity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1422):953–960, 1999.

[52] M. J. Keeling and K. T. Eames. Networks and epidemic models. *J. R. Soc. Interface*, 2(4):295–307, 2005.

[53] M. J. Keeling, T. D. Hollingsworth, and J. M. Read. The efficacy of contact tracing for the containment of the 2019 novel Coronavirus (COVID-19). *medRxiv*, 2020.

[54] H. Kim and A. Paul. Contact tracing: a game of big numbers in the time of COVID-19.

[55] I. Z. Kiss, D. M. Green, and R. R. Kao. Infectious disease control using contact tracing in random and scale-free networks. *J. R. Soc. Interface*, 3:55–62, 2006.

[56] I. Z. Kiss, D. M. Green, and R. R. Kao. The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *J. R. Soc. Interface*, pages 791–799, 2007.

[57] I. Z. Kiss, D. M. Green, and R. R. Kao. The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *J. R. Soc. Interface*, 5:791–799, 2008.

[58] D. Klinkenberg, C. Fraser, and H. Heesterbeek. The effectiveness of contact tracing in emerging epidemics. *PLoS ONE*, 1:e12, 2006.

[59] S. Kojaku, L. Hébert-Dufresne, and Y.-Y. Ahn. The effectiveness of contact tracing in heterogeneous networks.
[60] M. Kretzschmar, C. Satterwhite, J. Leichliter, and S. Berman. Effects of screening and partner notification on chlamydia positivity in the united states: a modeling study. *Sexually transmitted diseases*, 39(5):325–331, 2012.

[61] M. Kretzschmar, K. M. E. Turner, P. M. Barton, W. J. Edmunds, and N. Low. Predicting the population impact of chlamydia screening programmes: comparative mathematical modelling study. *Sexually Transmitted Infections*, 85(5):359–366, may 2009.

[62] M. Kretzschmar, S. Van den Hof, J. Wallinga, and J. Van Wijngaarden. Ring vaccination and smallpox control. *Emerging infectious diseases*, 10(5):832, 2004.

[63] M. Kretzschmar, Y. van Duynhoven, and A. Severijnen. Modeling prevention strategies for Gonorrhea and Chlamydia using stochastic network simulations. *Am J Epidemiol.*, 144(3):306–317, 1996.

[64] M. E. Kretzschmar, G. Rozhnova, M. Bootsma, M. E. van Boven, J. van de Wijgert, and M. Bonten. Time is of the essence: impact of delays on effectiveness of contact tracing for COVID-19. may 2020.

[65] M. E. Kretzschmar, G. Rozhnova, and M. E. van Boven. Isolation and contact tracing can tip the scale to containment of COVID-19 in populations with social distancing. *medRxiv*, 2020.

[66] A. J. Kucharski, P. Klepac, A. J. K. Conlan, S. M. Kissler, M. L. Tang, H. Fry, J. R. Gog, W. J. Edmunds, J. C. Emery, G. Medley, J. D. Munday, T. W. Russell, Q. J. Leclerc, C. Diamond, S. R. Procter, A. Gimma, F. Y. Sun, H. P. Gibbs, A. Rosello, K. van Zandvoort, S. Hué, S. R. Meakin, A. K. Deol, G. Knight, T. Jombart, A. M. Foss, N. I. Bosse, K. E. Atkins, B. J. Quilty, R. Lowe, K. Prem, S. Flasche, C. A. B. Pearson, R. M. J. Houben, E. S. Nightingale, A. Endo, D. C. Tully, Y. Liu, J. Villabona-Arenas, K. O’Reilly, S. Funk, R. M. Eggo, M. Jit, E. M. Rees, J. Hellewell, S. Clifford, C. I. Jarvis, S. Abbott, M. Azenbergs, N. G. Davies, and D. Simons. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases*, jun 2020.

[67] K. O. Kwok, A. Tang, V. W. Wei, W. H. Park, E. K. Yeoh, and S. Riley. Epidemic models of contact tracing: Systematic review of transmission studies of severe acute respiratory syndrome and middle east respiratory syndrome. *Comput. Struct. Biotechnol. J.*, 17:186–194, 2019.

[68] A. Lima, V. Pejovic, L. Rossi, M. Musolesi, and M. Gonzalez. Prognosis: Evaluating risky individual behavior during epidemics using mobile network data.

[69] F. Liu, W. T. A. Enanoria, J. Zipprich, S. Blumberg, K. Harriman, S. F. Ackley, W. D. Wheaton, J. L. Allpress, and T. C. Porco. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for california. *BMC Public Health*, 15(1), 2015.

[70] J. O. Lloyd-Smith, A. P. Galvani, and W. M. Getz. Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. *Proc. R. Soc. Lond. B Biol. Sci.*, 270(1528):1979–1989, 2003.

[71] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, nov 2005.
[72] L. Lorch, W. Trouleau, S. Tsirtsis, A. Szanto, B. Schölkopf, and M. Gomez-Rodriguez. A spatiotemporal epidemic model to quantify the effects of contact tracing, testing, and containment. arXiv, 2020.

[73] D. Lunz, G. Batt, and J. Ruess. To isolate, or not to isolate: a theoretical framework for disease control via contact tracing. may 2020.

[74] R. Mastrandrea and A. Barrat. How to estimate epidemic risk from incomplete contact diaries data? PLOS Computational Biology, 12(6):e1005002, jun 2016.

[75] G. A. McVean and N. J. Cardin. Approximating the coalescent with recombination. Philos. Transact. R. Soc. B, 360(1459):1387–1393, 2005.

[76] G. R. Mellor, C. S. M. Currie, and E. L. Corbett. Incorporating household structure into a discrete-event simulation model of tuberculosis and HIV. ACM Transact. Modeling Computer Simul., 21(4):1–17, 2011.

[77] Y. A. Melsew, M. Gambhir, A. C. Cheng, E. S. McBryde, J. T. Denholm, E. L. Tay, and J. M. Trauer. The role of super-spreading events in mycobacterium tuberculosis transmission: evidence from contact tracing. BMC Infe. Dis., 19(1), 2019.

[78] L. A. Meyers, B. Pourbohloul, M. Newman, D. M. Skowronska, and R. C. Brunham. Network theory and SARS: predicting outbreak diversity. J. Theor. Biol., 232(1):71–81, 2005.

[79] J. C. Miller, A. C. Slim, and E. M. Volz. Edge-based compartmental modelling for infectious disease spread. J. R. Soc. Interface, 9:890–906, 2011.

[80] K. Mizumoto, K. Ejima, T. Yamamoto, and H. Nishiura. Vaccination and clinical severity: Is the effectiveness of contact tracing and case isolation hampered by past vaccination? International Journal of Environmental Research and Public Health, 10(3):816, 2013.

[81] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Medicine, 5(3):e74, 2008.

[82] J. Müller and V. Hösel. Estimating the tracing probability from contact history at the onset of an epidemic. Math. pop. Stud., 14:211–236, 2007.

[83] J. Müller and B. Koopmann. The effect of delay on contact tracing. Math. Biosci., 282:204–214, 2016.

[84] J. Müller, M. Kretzschmar, and K. Dietz. Contact tracing in stochastic and deterministic epidemic models. Math. Biosci., 164:39 – 64, 2000.

[85] J. Müller and C. Kuttler. Methods and models in mathematical biology. Springer, 2015.

[86] J. Müller and M. Möhle. Family trees of continuous-time birth-and-death processes. J. Appl. Prob., 40:980–994, 2003.

[87] R. Naresh, A. Tripathi, and D. Sharma. A nonlinear HIV/AIDS model with contact tracing. Appl. Math. Comput., 217(23):9575–9591, 2011.
[88] M. E. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.

[89] A. Okolie and J. Müller. Exact and approximate formulas for contact tracing on random trees. *Math. Biosci.*, 321:108320, 2020.

[90] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi. Crowd sensing of traffic anomalies based on human mobility and social media. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - SIGSPATIAL 13*. ACM Press, 2013.

[91] D. K. Pasquale, I. A. Doherty, W. C. Miller, P. A. Leone, L. A. Sampson, S. L. Ledford, J. Sebastian, and A. M. Dennis. Factors associated with human immunodeficiency virus infections linked in genetic clusters but disconnected in partner tracing. *Sexually Transmitted Diseases*, 47(2):80–87, feb 2020.

[92] D. K. Pasquale, I. A. Doherty, L. A. Sampson, S. Hué, P. A. Leone, J. Sebastian, S. L. Ledford, J. J. Eron, W. C. Miller, and A. M. Dennis. Leveraging phylogenetics to understand HIV transmission and partner notification networks. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 78(4):367–375, aug 2018.

[93] C. M. Peak, L. M. Childs, Y. H. Grad, and C. O. Buckee. Comparing nonpharmaceutical interventions for containing emerging epidemics. *PNAS*, 114(15):4023–4028, 2017.

[94] C. M. Peak, R. Kahn, Y. H. Grad, L. M. Childs, R. Li, M. Lipsitch, and C. O. Buckee. Individual quarantine versus active monitoring of contacts for the mitigation of covid-19: a modelling study. *The Lancet Infectious Diseases*, 2020.

[95] T. C. Porco, K. A. Holbrook, S. E. Fernyak, D. L. Portnoy, R. Reiter, and T. J. Aragón. Logistics of community smallpox control through contact tracing and ring vaccination: a stochastic network model. *BMC Public Health*, 4(1), 2004.

[96] C. Rivers, E. Lofgren, M. Marathe, S. Eubank, and B. Lewis. Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia. *PLOS Currents Outbreaks*, Edition 2, 2014.

[97] A. Roetzer, R. Diel, T. A. Kohl, C. Rückert, U. Nübel, J. Blom, T. Wirth, S. Jaenicke, S. Schuback, S. Ritsch-Gerdes, P. Supply, J. Kalinowski, and S. Niemann. Whole genome sequencing versus traditional genotyping for investigation of a mycobacterium tuberculosis outbreak: A longitudinal molecular epidemiological study. *PLoS Medicine*, 10(2):e1001387, 2013.

[98] M. Salathe, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, 107(51):22020–22025, dec 2010.

[99] K. Sato, H. Matsuda, and A. Sasaki. Pathogen invasion and host extinction in lattice structured populations. *J Math Biol*, 32:251–268, 1994.

[100] N. Shaban, M. Andersson, Åke Svensson, and T. Britton. Networks, epidemics and vaccination through contact tracing. *Mathematical Biosciences*, 216:1–8, 2008.

[101] N. M. Shahtori, T. Ferdousi, C. Scoglio, and F. D. Sahneh. Quantifying the impact of early-stage contact tracing on controlling ebola diffusion. *Math Biosc Engin*, 15(5):1165–1180, 2018.
[102] K. J. Sharkey, I. Z. Kiss, R. R. Wilkinson, and P. L. Simon. Exact equations for SIR epidemics on tree graphs. *Bull. Math. Biol.*, 77:614–645, 2015.

[103] J. Stehlé, N. Voisin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Régis, J.-F. Pinton, N. Khanafer, W. Van den Broeck, et al. Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. *BMC medicine*, 9(1):87, 2011.

[104] T. Tanaka, T. Yamaguchi, and Y. Sakamoto. Estimation of the percentages of asymptomatic patients and undiagnosed patients of the novel coronavirus (SARS-CoV-2) infection in Hokkaido, Japan by using birth-death process with recursive full tracing. *medRxiv*, 2020.

[105] L. Taylor. No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environ. Plan. D: Soc. Space*, 34(2):319–336, 2015.

[106] Y. Tian, F. Alawami, A. Al-Azem, N. Osgood, V. Hoeppner, and C. Dutchyn. A system dynamics model of tuberculosis diffusion with respect to contact tracing investigation. In *Proceedings of the 2011 Winter Simulation Conference (WSC)*. IEEE, dec 2011.

[107] Y. Tian, N. D. Osgood, A. Al-Azem, and V. H. Hoeppner. Evaluating the effectiveness of contact tracing on tuberculosis outcomes in saskatchewan using individual-based modeling. *Health Educ. Behav.*, 40(1_suppl):98S–110S, 2013.

[108] L. S. Tsimring and R. Huerta. Modeling of contact tracing in social networks. *Physica A*, 325(1-2):33–39, 2003.

[109] K. Turner, E. Adams, N. Gay, A. Ghani, C. Mercer, and W. J. Edmunds. *Theoretical Biology and Medical Modelling*, 3(1):3, 2006.

[110] R. R. Wilkinson and K. J. Sharkey. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Phys. Rev. E*, 89:022808, 2014.

[111] L. Willem, F. Verelst, J. Bilcke, N. Hens, and P. Beutels. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infect. Dis.*, 17(1), 2017.
## A Literature overview

The table orders the papers primarily according to the infection they aim to investigate resp. for the rather theoretical papers, by the method used.

| author                  | infection                                     | method            | term for CT                          | CT delay | remarks                                                                 | outcome(w.r.t. CT)                                                                 |
|-------------------------|-----------------------------------------------|-------------------|--------------------------------------|----------|------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Becker et al. 2005 [9]  | SARS                                          | phenom. /         | reduction of incidence               |          | article exclusively aims at Reff, no dynamics.                        | formula for Reff; social distancing together with CT can control SARS            |
|                         |                                               | next generation   | operator                              |          | based on Fraser [30]; only Reff, no dynamics; airborne infection      |                                                                                   |
| Chen et al. 2006 [16]   | Influenza, Measles, Chickenpox, SARS          | PDE               | fixed fraction of newly infected      | -        | review article                                                         | Probability to control an outbreak is estimated                                   |
|                         |                                               |                   | are eventually traced                 |          |                                                                        |                                                                                   |
| Kwok et al. 2019 [67]   | SARS                                          |                   |                                      |          |                                                                        |                                                                                   |
| Fraser et al. 2004 [30] | Theory & Influenza, SARS, Smallpox, HIV       | PDE               | fixed fraction of newly infected      | -        | basic model, used in applications [16, 30]; timing of incubation period and latency period is central for CT |                                                                                   |
| Lloyd-Smith et al. 2003 | SARS                                          | phenom. /         | increased rate to quarantine          | +        | CT not explicitly formulated                                          | crucial that CT is implemented at the beginning of the outbreak                 |
|                         |                                               | time-discrete     |                                      |          |                                                                        |                                                                                   |
|                         |                                               | stoch. sim on pop.-level |                        |          |                                                                        |                                                                                   |
| Bradshaw et al. 2020 [13]| SARS-CoV-2                                    | IBM, homogeneous  | direct formulation                   | +        | preprint; DCT considered, focus on onset, Reff and prob. for major outbreak | backward tracing and high abundance of DCT devices necessary                     |
|                         |                                               | population        |                                      |          |                                                                        |                                                                                   |
| Bulchandani et al. 2020 | SARS-CoV-2                                    | IBM, homogeneous  | direct formulation                   | -        | preprint; DCT considered, focus on onset, heuristic formula for Reff   | High DCT-device coverage necessary                                               |
|                         |                                               | population        |                                      |          |                                                                        |                                                                                   |
| Ferretti et al. 2020    | SARS-CoV-2                                    | phenom. /         | fixed fraction of newly infected      | +        | based on Fraser [30]; DCT considered, heuristic formula for Reff       | SARS-CoV-2 controllable by DCT                                                   |
|                         |                                               | PDE               | are eventually traced                 |          |                                                                        |                                                                                   |
| Giordano et al. 2020    | SARS-CoV-2                                    | ODE/SIR           | increased rate for removal            | -        | detailed ODE model                                                     | CT is a central element in controlling the infection                            |
|                         |                                               |                   |                                      |          |                                                                        |                                                                                   |
| Helleswell et al. 2020  | SARS-CoV-2                                    | IBM, homogeneous  | direct formulation                   | -        | onset of epidemic considered                                          | high tracing probability necessary to control the infection                      |
|                         |                                               | population        |                                      |          |                                                                        |                                                                                   |
| Hernandez-Orallo et al. | SARS-CoV-2                                    | IBM, homogeneous  | direct formulation                   | +        | simulations based on empirical contact network                        | CT needs to be precise in order or avoid many persons in quarantine              |
|                         |                                               | population        | (IBM)                                 |          |                                                                        |                                                                                   |
| author                | infection     | method                        | term for CT | CT delay | remarks                                                                 | outcome(w.r.t. CT) |
|-----------------------|---------------|-------------------------------|-------------|----------|--------------------------------------------------------------------------|-------------------|
| Keeling et al. 2020   | SARS-CoV-2    | IBM, homogeneous population  | direct formulation |         | preprint; time and intensity of contacts vary, estimation of $R_{eff}$   | high tracing probability necessary to control the infection |
| Kim et al. 2020       | SARS-CoV-2    | Stochastic branching process  | verbally     |         | preprint; DCT considered; Heuristic calculations of efficiency           | heuristic formula for efficiency |
| Kretzschmar et al. 2020| SARS-CoV-2    | IBM, homogeneous population  | direct formulation           |         | preprint; Model based on Kretzschmar(2004)                                | middle range tracing probability necessary to control diseases |
| Kretzschmar et al. 2020| SARS-CoV-2    | IBM, homogeneous population  | direct formulation           |         | preprint; Model based on Kretzschmar [62]; DCT considered, combined with conventional CT | DCT considered, combined with conventional CT |
| Kucharski et al. 2020  | SARS-CoV-2    | IBM, homogeneous population  | direct formulation           |         | preprint; DCT considered, combined with conventional CT                  | CT more efficient than mass testing |
| Lorch et al. 2020      | SARS-CoV-2    | IBM with discrete spatial structure | one-step tracing, if at similar times in the same location |         | DCT considered, combined with conventional CT                           | DCT efficient particularly in case of a low fraction of quarantined persons |
| Lunz et al. 2020       | SARS-CoV-2    | phenomen. / ODE SEIR          | mass action                |         | preprint; the tracing rate is connected to contact heterogeneity but not based on first principles | optimal CT defined as minimizing the total number of individuals that go into quarantine during the outbreak |
| Tanaka et al. 2020     | SARS-CoV-2    | IBM, homogeneous population  | direct formulation           |         | simulation of clusters detected by CT as input for stats (see also Blum [12]) | Bayesian parameter estimation based on CT |
| Berge et al. 2018      | Ebola         | phenomen. / ODE SEIR          | fixed fraction of newly infecteds go to quarantine |         | CT not explicitly formulated                                             | stationary states and their stability analyzed |
| Browne et al. 2015     | Ebola         | branching process/ODE, SEIR   | fixed fraction of newly infecteds go to quarantine |         | fraction of detected cases is computed based on the branching-process    | paper aims at a theoretical framework that is feasible for practical applications |
| Shahtori et al. 2018   | Ebola         | IBM, homogeneous population  | direct formulation           |         | onset of infection                                                       | crucial that CT is implemented at the beginning of the outbreak |
| Rivers et al. 2014     | Ebola         | phenomen. / stoch. sim on pop.-level./ODE, SEIR | increased diagnosis rate |         | CT not explicitly formulated                                             | reduction of $R_{eff}$ by around 30% possible |
| author        | infection | method | term for CT | CT delay | remarks                                                                 | outcome (w.r.t. CT) |
|--------------|-----------|--------|-------------|----------|-------------------------------------------------------------------------|---------------------|
| de Araozoa et al. 2002 [21] | HIV | phenomen. / mass action | ODE SI | - | three classes of infecteds modeled; not know their infection, AIDS       | stability analysis of stationary points, comparison with data |
| Clémerançon et al. 2008 [20] | HIV | phenomen. several / stoch. terms: linear, mass action, saturation | ODE, SI | - | extension of de Azoza [21]                                              | development of statistical tools (maximum likelihood estimators) |
| Blum et al. 2010 [12] | HIV | phenomen. linear and saturation / stochastic model on population level | + | see also [59, 101] | Bayes interference (ABC and Metropolis Hastings) for estimating CT probability |
| Hsieh et al. 2005 [22] | HIV | phenomen. / several terms: linear, mass action, saturation | ODE SI | - | extension of de Azoza [21]                                              | mass action inappropriate, linear or saturation term for CT better |
| Hsieh et al. 2010 [36] | HIV | phenomen. / function: two-step tracing | ODE SI | - | stability analysis of stationary points, Reff; two-step tracing superior over one-step tracing |
| Hyman et al. 2003 [15] | HIV | phenomen. / mass action | ODE SI | - | two models: core group, different stages of HIV                           | Reff, sensitivity analysis |
| Mellor et al. 2001 [76] | HIV and Tuberculosis | IBM with screening the household structure | - | cross-tracing of HIV and Tuberculosis is effective |
| Naresh et al. 2006 [87] | HIV | phenomen. / fixed fraction of newly infecteds know their infection | ODE SI | - | stability analysis of stationary points                                 | CT is not triggered by diagnosis, but infections of known infecteds |
| Clark et al. 2012 [18] | Chlamydia | pair approx. removal rate on infected-diagnosed pairs | - | based on [25, 40] | CT is efficient particularly below a certain prevalence                  |
| Clark et al. 2013 [19] | Chlamydia | IBM + IBM: direct; ODE: power law in I | - | power law adapted to IBM simulations; optimal resource allocations (CT/screening) |
| Heffernan et al. 2009 [36] | Chlamydia | phenomen. mass action | ODE SI | - | model includes random screening (yield index cases) and CT               | model results in line with data |
| Hethcote et al. 1984 [39] | Gonorrhea | phenomen. decreased incidence | ODE | - | introduce backward and forward tracing                                    | pioneering work about CT |
| Kretzschmar et al. 1996 [7] | Gonorrhea, Chlamydia | IBM with household structure of a fraction p of partners | - | one-step tracing                                                        | prevalence for different control scenarios |
| Hline | | | | | | |
| author               | infection | method          | term for CT | CT delay | remarks                                                                 | outcome(w.r.t. CT) |
|----------------------|-----------|-----------------|-------------|----------|--------------------------------------------------------------------------|--------------------|
| Kretzschmar et al. 2009 [61] | Chlamydia | IBM direct formulation | +           | three different IBM models previously published | due to their complexity | the results of the models are somewhat different, compared |
| Turner et al. 2006 [107] | Chlamydia | IBM direct formulation | +           | model with pair formation, CT within pairs | Model fits data, and yields comparable results comparable studies |
| Aparicio et al. 2006 [3] | Tuberculosis | phenomen. / ODE SEIR | fixed fraction of newly infecteds are identified | CT not explicitly formulated | simulation of prevalence |
| Begun et al. 2013 [10] | Tuberculosis | IBM with screening the household structure | household tracing reduces the incidence by 2%-3% | review article |
| Kassaie et al. 2014 [19] | Tuberculosis | IBM with household structure | - contacts outside the household are not traced | simulation of prevalence |
| Tian et al. 2011 [105] | Tuberculosis | IBM, inhomogeneous population direct formulation | different scenarios, sensitivity analysis | simulation of prevalence |
| Tian et al. 2012 [10] | Tuberculosis | IBM, inhomogeneous population direct formulation | Follow-up of Tien [106] | simulation of prevalence |
| Agarwal et al. 2012 [1] | Influenza | phenomen. / ODE SIR | fraction of newly infecteds go to quarantine | two risk classes in susceptibles are considered | Dynamical systems analysis, Reff |
| Eichner 2003 [27] | Smallpox | phenomen. / Stochastic branching process | all close contacts and fraction of contacts are traced | critical tracing probability estimated |
| Kaplan et al. 2002 [47] | Smallpox | phenomen. / ODE function | contactees who are traced are vaccinated | mass vaccination superior to vaccination triggered by CT |
| Kretzschmar et al. 2003 [29] | Smallpox | Stochastic branching process | ring vaccination triggered by contact tracing | delay in CT is crucial |
| Porco et al. 2004 [95] | Smallpox | IBM with household structure direct formulation | one step and two-step tracing compared | massive CT and ring vaccination can control the outbreak |
| Liu et al. 2015 [60] | Measles | IBM, inhomogeneous population direct formulation | complex/realistic contact structure | CT can significantly contribute to control a measles outbreak |
| Ball et al. 2011 [7] | Theory | Stochastic branching process direct formulation | SIR, focus on fixed and exponentially distributed infectious period | analytic approach, bounds on Reff, extinction probability |

21
| author            | infection            | method                                      | term for CT                   | CT delay, remarks                                                                 | outcome (w.r.t. CT)                                                                 |
|-------------------|----------------------|---------------------------------------------|-------------------------------|---------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Ball et al. 2015  | Theory               | Stochastic branching process                | direct formulation           | + SEIR, follow up of Ball [7]                                                   | effect of tracing delay, Reff, extinction probability                             |
| Müller et al. 2000| Theory               | Stochastic branching process                | direct formulation           | - focus on age since infection                                                   | Reff, ODE approximation, critical tracing probability                             |
| Müller et al. 2007| Theory               | Stochastic branching process                | direct formulation           | - based on Müller [84] see also [12, 104]                                       | estimation of tracing probability                                                 |
| Müller et al. 2016| Theory               | Stochastic branching process                | direct formulation           | + based on Müller [84]                                                          | effect of tracing delay and latency period, Reff                                   |
| Okolie et al. 2018| Theory               | Stochastic branching process                | direct formulation           | - connects branching process and pair approx., based on Müller [84]             | effect of a random contact graph on CT                                             |
| Klinkenberg et al.2006 | Theory & influenza, smallpox, SARS, and foot-and-mouth disease | Stochastic branching process | direct formulation           | + single-step and recursive tracing                                             | mostly: single step and recursive tracing is equal effective                      |
| Shaban et al. 2008| Theory               | Stochastic branching process                | direct formulation           | - Vaccination of detected individuals                                             | Reff, probability for extinction, simulation of final size                        |
| Kojaku et al. 2020| Theory               | Stochastic branching process                | direct formulation           | - CT on random graph, focus on generating functions for the degree              | CT highly efficient as nodes with high degree are traced                          |
| Kiss et al. 2006  | Theory               | IBM, inhomogeneous population              | direct formulation           | - isolation of susceptible; scale free and Poisson network                      | For scale free networks, tracing effect less sensitive to the epidemiological parameters |
| Kiss et al. 2008  | Theory               | IBM, inhomogeneous population              | direct formulation           | - assortively / disassortatively mixing networks; Single-step and recursive tracing | CT more effective in disassortative networks; recursive tracing more efficient     |
| Farrahi et al. 2008| Theory               | IBM, inhomogeneous population              | direct formulation           | + first model for digital CT                                                    | digital CT can be efficient even if the fraction of app-users is small            |
| Eames et al. 2002 | Theory / STD         | pair approximation & stochastic simulation  | direct formulation           | - this paper introduced pair approximation for CT based on Eames [25]          | modeling CT by pair approximation                                                 |
| Eames et al. 2003 | Theory               | pair approximation & stochastic simulation  | direct formulation           | - based on Eames [25]                                                          | critical tracing probability                                                     |
| Eames et al. 2005 | Theory               | pair approximation & stochastic simulation  | direct formulation           | - based on [25]; focus on different social graphs (small world, scale-free)     | network structure influence efficiency of CT                                      |
| author          | infection | method | term for CT | CT delay | remarks | outcome(w.r.t. CT)                      |
|-----------------|-----------|--------|-------------|----------|---------|----------------------------------------|
| Eames 2007 [24]| Theory    | pair approximation & stochastic simulation | direct formulation | -        | based on Eames [25] | recursive CT much more effective than one-step CT; “targeted CT”; focus on core groups |
| House et al. 2010 [10]| Theory | pair approximation & stochastic simulation | direct formulation | -        | based on Eames [25]; focus on different social graphs (small world, scale-free) | CT has higher efficiency in clustered contact graphs |
| Huerta et al. 2002 [14]| Theory | pair approximation & stochastic simulation | direct formulation | -        | develop pair approximation for CT | rewiring of contact network decreases CT |
| Tsimering et al. 2003 [118]| Theory | pair approximation & stochastic simulation | direct formulation | -        | based on Huerta [14] | rewiring of contact network decreases CT |
| Armbruster et al. 2007 [4] | Theory | phenomenological / ODE | linear removal term | -        | addresses a cost-efficiency analysis based on Armbruster [4] | CT only cost efficient if prevalence is small |
| Armbruster et al. 2007 [5] | Theory | phenomenological / ODE | linear removal term | -        | focus on the onset; multitype branching process, analyzed by generating functions | CT only cost efficient if prevalence is small |
| Mizumoto et al. 2013 [80] | Theory | phenomenological / next generation operator | reduction of R0 by a factor | -        | Reff, probability for extinction, duration of a minor outbreak | Reff, probability for extinction, duration of a minor outbreak |