Boyd, H. A., Laksafoss, A. D., Hjalgrim, H., & Poulsen, G. J. (2026). Applying the methods of chronic disease epidemiology to the study of compassion. *International Journal of Wellbeing*, *16*(2), 5565, 1-21. https://doi.org/10.5502/ijw.v16i2.5565

ARTICLE

Applying the methods of chronic disease epidemiology to the study of compassion

Heather A. Boyd · Anna Damkjær Laksafoss · Henrik Hjalgrim
Gry Juul Poulsen

Abstract: Now, more than ever, the world needs compassion—at the level of the individual, the community, the workplace, and society as a whole. As an abstract concept, the idea that increasing compassion will benefit society on many levels is simple. But as a target for public health intervention, our fundamental understanding of compassion is poor. How can we craft interventions to increase our willingness and capacity to help others through adversity and suffering when our understanding of the factors that increase compassion, the barriers to compassion, and their interplay, is limited? Work by a wide range of social and behavioral scientists has laid the foundations for compassion studies by establishing consensus definitions and identifying recurring themes in compassionate interactions. We believe that applying epidemiologic methods to study the determinants of compassion will provide a complementary, quantitative foundation for future initiatives, interventions, and benchmarking. In this paper, we explore how compassion has many of the same features as chronic diseases, including multiple dimensions, time scales, and degrees of modifiability, evolution over time, non-linear progression, and complex, interacting influences. We weigh the suitability of various epidemiologic study designs for studying compassion. We then propose a number of analytic approaches from chronic disease epidemiology that could be used to build quantitative, and potentially even causal, models of compassion. We examine the advantages that would be gained by using methods such as multi-level modeling, sequential conditional mean modeling, marginal structural modeling, and machine learning to study compassion and its influences, along with the potential limitations of each method. Finally, we round out the paper with a discussion of the challenges involved in using epidemiologic frameworks to study compassion.

Keywords: chronic disease; compassion; epidemiology; statistical modelling; study design

1. Introduction

The idea that increasing the amount of compassion in a society will improve the wellbeing of its people is simple in the abstract. Current world events large and small speak to the terrible consequences of a lack of compassion at individual, organizational, community, regional, national, and international levels. A sizeable body of literature has identified recurring themes in compassionate interactions, and behavioral interventions designed to increase compassion have been tested at individual, institutional, and regional levels (see for example the state-wide compassionate leadership intervention trial in hospitals in Bihar, India, described by Jha et al. (2026). However, it is often difficult to quantify the impact of giving or receiving compassion,

Heather A. Boyd
Statens Serum Institut, Denmark
hoy@ssi.dk

even at the individual level, and in practice, tailoring an intervention to a specific population can be challenging, as our fundamental understanding of what compassion is and how it "works" under different conditions remains rudimentary.

Central to any study of compassion is the challenge of defining what constitutes compassion, a task greatly influenced by religious and cultural traditions and scholarly disciplines. Depending on the lens through which it is viewed, compassion is variously considered an emotional, active response to suffering, a motivational state with physiological underpinnings, a social practice shaped by institutions and culture, a moral virtue, and a spiritual principle (Addiss et al., 2022). However, three features are common to most definitions: 1) awareness of suffering; 2) empathy in the face of suffering; and 3) action or the intention to act to relieve or prevent suffering.

A second challenge in studying how to promote compassion is determining how to quantify current capacity for compassion and track changes in this capacity over time. In the Buddhist tradition, compassion is considered unmeasurable and unquantifiable; at any given moment, an individual does or does not display compassion, and persons perceived to be particularly compassionate simply demonstrate more compassionate moments than others (Salzberg, 1995). However, in recent years researchers have applied methods from the social and behavioral sciences in early attempts to define and quantify compassion (e.g., Strauss et al., 2016; Gu et al., 2020).

Ozawa-de Silva and Mascaro delve more deeply into the challenges associated with defining compassion in a companion paper in this issue (Ozawa-de Silva & Mascaro, 2026). Assuming that clearly-defined measures of the capacity for compassion can be arrived at, we believe that applying epidemiologic methods to study compassion will complement previous qualitative studies. In simple terms, epidemiologic studies describe and quantify the relationship between an exposure and an outcome of interest in the context of other factors that might influence this relationship, either by explaining all or part of an observed relationship between exposure and outcome (confounding variables) or by modifying this relationship (interaction with effect modifiers). Epidemiologic studies of compassion as an outcome will allow us to quantify the degree to which various factors increase, or serve as barriers to, the capacity for compassion and to understand how these factors interact. Focusing studies on compassion as an exposure will allow us to determine how giving or receiving compassion affects a wide variety of outcomes experienced by the giver, the receiver, and their communities. Finally, intervention studies will allow us to evaluate the effectiveness of deliberate measures to increase compassion. In this paper, we focus on the use of epidemiologic methods to study compassion as an outcome, based on the position that a solid, quantitative understanding of the determinants of compassion is necessary before we can determine how and where to intervene to increase compassion or understand the downstream effects of compassion.

The science of epidemiology is rooted in a fundamental focus on person, place, and time, on necessary and sufficient causes, and on the influences of confounding and bias. Its methods are therefore well suited to unravelling the complex combinations of short- and long-term influences and cumulative exposures over time that likely govern expressions of compassion. They can also be tailored to account for elapsed time, timing of events, and context when determining the absolute and relative impacts of individual determinants of compassion.

When choosing methods to study the determinants of a given disease, the epidemiologist must first identify disease features that the method must be able to accommodate. Is the disease transmissible? If so, what is the route of transmission? Does survival guarantee future immunity or is a return to the susceptible state possible after acute illness? Is the disease progressive and if

so, does it always proceed according to the same pattern or are multiple patterns of progression possible? Is it steadily progressive or is remission possible? Is it possible that the disease could have multiple etiologies and develop in multiple ways or that multiple "hits" might be necessary to produce pathology? Are genetic influences monogenic or polygenic?

A variety of epidemiologic methods might contribute usefully to the study of particular features of compassion. For example, the capacity for compassion resembles predisposition to infectious disease both in its propensity for repeated periods of susceptibility and "immunity" over time and its transmissibility. Methods frequently used to study infectious disease transmission might prove useful to investigate these features of the capacity for compassion. In the current paper, however, we discuss what the methods used in chronic disease epidemiology could contribute to the study of compassion.

2. Treating compassion as a chronic disease: an epidemiologist's perspective

From an epidemiologist's point of view, compassion shares a number of features with chronic diseases such as ischemic heart disease (IHD). Known risk factors for IHD (Brown et al., 2023) and factors thought to play a role in increasing or decreasing the capacity for compassion (Addiss et al., 2022) span multiple dimensions. In both cases, these factors range from personal traits and personal circumstances to interpersonal and community-level factors and structural and environmental factors, and exhibit varying degrees of time dependence and modifiability (Figure 1).

In addition, neither IHD (or its precursors) nor the capacity for compassion are static conditions. Both evolve over time in response to positive or negative short- and long-term influences and life events (Figure 2). Furthermore, that evolution does not necessarily progress in a linear fashion. An individual who is genetically predisposed to IHD and has worrisome laboratory results may, through dint of lifestyle changes and medical intervention, never develop symptomatic disease, while the first inkling of a problem in a person with no obvious risk factors may come in the form of a massive myocardial infarction. Similarly, the capacity for compassion may fluctuate up and down, depending on positive and negative pressures and life events.

Finally, the degree to which certain factors are associated with the risk of IHD often depends on the presence or absence of other risk factors (Li et al., 2023). Likewise, early forays into the epidemiology of compassion suggested that the various potential determinants of compassion may interact in complex ways and involve the responses of both the provider and the recipient of compassion (Addiss et al., 2022). For example, volunteerism and community service are generally associated with increased levels of compassion in the provider, but if such service is required (e.g., to achieve school credit or as mandated by a legal judgment), providers often exhibit decreased levels of compassion (Lovette-Colver, 2013). As the number of persons affected by an accident or a natural disaster increases, levels of compassion among those responding or considering a response to the event have been shown to diminish (Butts et al., 2019), unless those providing compassion have themselves experienced adversity and perceive themselves as well-equipped to help, in which case their capacity for compassion increases (Lim & DeSteno, 2020). Similarly, while increasing distress in the compassion recipient is associated with more compassionate responses in compassion providers who perceive themselves as having low social power, compassion providers with high social power respond to increasing distress in the recipient with less compassionate responses (van Kleef et al., 2008).

PERSONAL TRAITS	PERSONAL CIRCUMSTANCE	COMMUNITY	STRUCTURAL/ ENVIRONMENTAL
Childhood SES	Age	Interpersonal relationships	Inadequate public transportation
Education and medical literacy	Current SES, reversal, unemployment	Loneliness	Institutional racism/structural stress
Genetics	IHD risk factors (e.g., hypertension, obesity)	Marital status	Non-walkable neighbourhoods
Lifestyle (smoking, physical activity, diet)	Menopause	Social contacts	Poverty
Race/ethnicity	Psychosocial stress	Support services	Unavailability of fresh produce
Sex	Other disease (e.g. diabetes, kidney disease)		

Range from fixed and invariant to time-varying and highly modifiable	Time-varying, variety of time scales and degrees of reversibility, not always modifiable	Time-varying, medium to long time scales, frequently not modifiable	Time-varying but only slowly and over a long period of time, difficult to modify

PERSONAL TRAITS	PERSONAL CIRCUMSTANCE	COMMUNITY	STRUCTURAL/ ENVIRONMENTAL
Attachment	Age	Church attendance	Legal frameworks
Emotional and social intelligence	Expectation, time pressure	Community service, volunteerism	Political leadership
Openness, humility	Relational closeness	In-group membership	Risks and rewards
Sex, race	Social power	Social norms	Societal norms
Religious faith	Stressful life events	Workplace leadership	Tradition
Perspective-taking, empathy	Target distress, severity of suffering, numbers		
	Well-being, physical and mental health		

Figure 1. Multidimensionality of known risk factors for ischemic heart disease (top) and factors thought to influence the capacity for compassion (bottom). In both cases, risk factors fall into one of four broadly-defined categories and exhibit varying degrees of time dependence and modifiability. IHD, ischemic heart disease. SES, socioeconomic status.

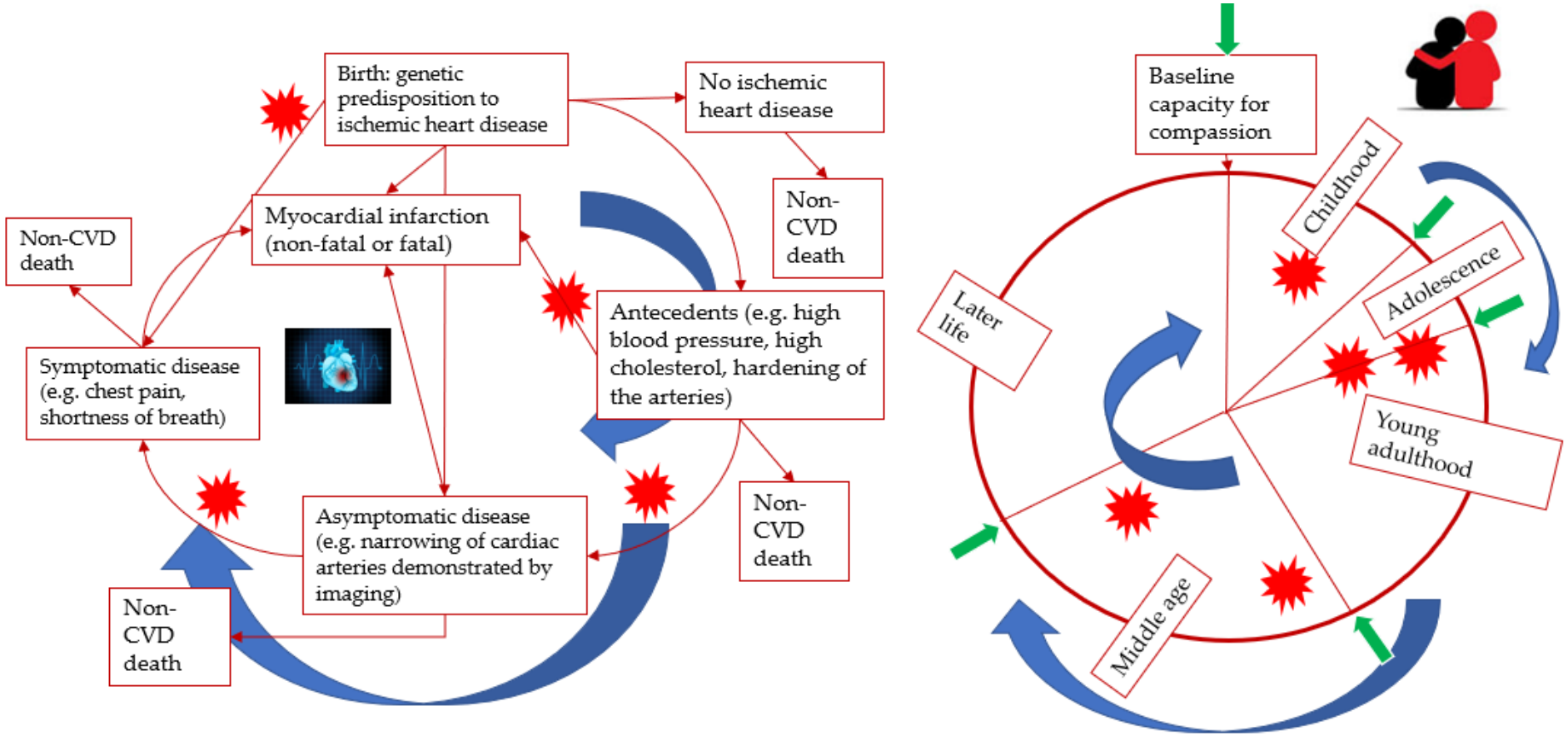

Figure 2. Illustration of the evolving and potentially non-linear nature of ischemic heart disease (left) and capacity for compassion (right) from birth over the life course, along with the timing of their respective short- and long-term determinants. Red explosions: time-limited life events and experiences. Blue arrows: long-term influences such as lifestyle factors (e.g. smoking and exercise) or medical interventions (e.g. blood pressure medication and stenting of cardiac arteries) for ischemic heart disease, and poverty or church attendance for compassion. Green arrows: times at which capacity for compassion is assessed in a hypothetical study. CVD, cardiovascular disease.

3. Study designs previously used to study compassion: strengths and limitations

Addiss and colleagues recently reviewed existing studies that applied epidemiologic methods to investigate determinants of compassion and found that the study designs used generally fell into one of four categories: cross-sectional designs (39.0% of studies), randomized experimental studies (30.5%), non-randomized or uncontrolled experimental designs (17.1%), and longitudinal cohort studies (8.5%) (Addiss et al., 2022). In this section, we briefly review the key features and general strengths and limitations of these four study designs and highlight important issues to consider when analyzing data from each type of study. We also offer general comments on design features that lend themselves to compassion studies and features that limit the ability of a given study type to address questions relating to compassion.

3.1 Cross-sectional studies

Cross-sectional studies are observational studies that assess both exposure(s) (in chronic disease epidemiology, generally risk factors or disease determinants) and outcome(s) (aspects of the disease of interest) together at a single point in time, producing a 'snapshot' of the study population at that moment. Because all study variables are assessed simultaneously, there is no follow-up of study participants and no 'drop-out' of the kind that otherwise results when time passes between the ascertainment of exposure and outcome or when measurements are repeated on multiple occasions over time. A well-conducted cross-sectional study can therefore collect an enormous amount of information on a large study population within a relatively short period of time. However, collecting information at a single point in time can make it challenging to establish a timeline for each factor relative to the others, particularly if a long time has elapsed since the events a study participant is being asked to recall. As a result, although cross-sectional studies can quantify associations between potential exposures and outcomes of interest, the inability to establish temporality (i.e., to establish that factor X preceded factor Y) means that cross-sectional studies cannot be used to demonstrate causality. Published cross-sectional studies of compassion were typically questionnaire-based survey studies, which had the advantage of being able to collect a great deal of information on a wide variety of factors potentially related to compassion. On the other hand, while these studies produced a great deal of information about potential determinants of interest, the inability to establish with certainty that determinant X preceded compassion level Y, rather than the reverse, limited the studies' ability to draw conclusions about causal relationships. Furthermore, assessing capacity for compassion at a single moment in time and then relating this measure to covariate levels summarized over a study participant's entire history also assumes that the capacity for compassion is invariant over time or cumulative over the lifespan to date, when it might actually be episodic or demonstrate other patterns.

3.2 Randomized experimental studies

In randomized experimental studies, researchers apply the randomized controlled trial (RCT) design that is the gold standard in pharmaceutical industry tests of new medications and is the only study design that can directly establish causality. When used to evaluate the efficacy and safety of a new medication, eligible study participants are randomly assigned treatment with either the new medication or standard treatment (or placebo, if there is no existing treatment) and then observed over a prespecified period of time to determine whether rates of predefined outcomes and side effects differ between the two groups. The advantage of this study design is that when done correctly, only treatment differs between the two groups. By assigning study

participants to treatment groups completely at random, all other variables that could affect the study outcomes should be similarly distributed in the two groups, an important RCT design feature that minimizes bias due to confounding by both known and unknown confounders. (Confounding results when a factor that is not the treatment of interest, the confounder, both is associated with treatment [e.g., by being more prevalent in one of the treatment groups] *and* affects the outcome, with the result that part or all of the observed association between the treatment and the outcome is actually produced by the confounding variable.) Furthermore, in the double-blinded version of the RCT design, neither the study participant nor the researcher tasked with assessing well-defined outcomes knows which treatment group the study participant is assigned to, minimizing surveillance bias (bias that arises when one treatment group is observed more closely for the outcomes of interest than the other group). Disadvantages associated with RCTs relate primarily to the rigidity of treatment group assignment, which ignores potential side effects, other reasons for non-adherence or partial adherence to treatment, and general loss to follow-up. This discordance has necessitated the development of statistical methods to accommodate deviation from assigned treatment protocol. In addition, RCTs are specifically designed around prespecified target outcomes, but because they are frequently resource-intensive to conduct, there is typically great interest in using their data for secondary analyses. Performing such 'post-hoc' analyses may require additional attention to the potential for residual confounding and inadequate power to examine associations with the secondary outcomes. Randomized experimental studies are ideal for studying interventions to increase compassion. However, rigorous maintenance of blinding and unbiased outcome assessment may be difficult when evaluating the impact of event scenarios on behavioural outcomes like capacity for compassion. Furthermore, in contrast to studies of disease endpoints such as myocardial infarction or cancer survival, compassion study participants can identify desirable responses and directly influence their own endpoints, potentially even acting in opposition to their natural inclination.

3.3 Non-randomized or quasi-experimental studies

Non-randomized or quasi-experimental studies share many of the features of RCTs, with one notable difference: treatment/intervention group assignment is not random and may depend on clinical considerations, patient choice, or other factors that may confound the results if not accounted for in the analysis. The so-called 'non-RCT' design may be useful in situations where ethical constraints prevent randomization or the goal is to test the efficacy of an intervention under real-world conditions (as opposed to in the somewhat artificial, rigidly-controlled setting required by an RCT). However, because they allow the distribution of potential confounding variables to vary across treatment groups, they more closely resemble observational studies in that control for confounding is limited to known confounders that are measured and adequately adjusted for in the analyses. Consequently, they are more subject to bias than RCTs and their results cannot necessarily be interpreted as evidence of causation.

3.4 Longitudinal cohort studies

Longitudinal cohort studies are a specific type of observational cohort study in which study participants are followed for the outcome(s) of interest over a long period of time, with repeated contacts with research staff to assess changes in exposure status and measure other potentially important variables that vary over time. The advantages of this study type include the ability to examine timing, patterns, and cumulative effects of exposures and other covariates that change over time and a reduction in recall bias relative to the cross-sectional study. Drawbacks include

the financial and logistical challenges associated with conducting longitudinal studies and the need for complex statistical methods in the analysis phase. To implement this design effectively, researchers need to carefully consider the timelines on which the factors to be studied likely operate and allow for adequate follow-up time; the longer the follow-up period and the more frequent the contact with study participants, the more resources are necessary to conduct the study and the greater the chance that participants will be lost to follow-up. Analyzing the data from longitudinal studies requires the use of statistical methods that can accommodate time-varying variables, as the levels of exposures, confounders, effect modifiers (covariates that interact with the exposure of interest, such that the exposure-outcome association differs according to the level of the effect modifier), and even outcomes can have their own temporal patterns over the study period. Furthermore, statistical methods used to analyze longitudinal data must also be able to account for the correlation between repeated measurements made on the same individual (two measurements from the same person at different times will be more similar than two measurements from two different people).

To date, longitudinal cohort studies have been least used in the study of compassion, although they are arguably the non-experimental design most fit for this purpose. Through repeated contact with study participants over time, this design provides a framework for studying how compassionate responses change over time in response to changes in potential determinants. Consequently, in the following section, we focus primarily on statistical methods for analyzing data from longitudinal cohort studies.

4. Statistical methods from chronic disease epidemiology potentially suitable for modelling compassion as the outcome in an epidemiologic study

In analyzing data from longitudinal cohort studies, chronic disease epidemiology employs a range of statistical methods, many of which might also be useful in epidemiologic studies of compassion. These methods fall into three broad groups: multi-levels models, methods that accommodate time-varying covariates, and machine learning methods. For simplicity, we describe the groups as separate entities with specific strengths and weaknesses, but the groups overlap somewhat, with some methods falling in more than one group.

4.1 Multi-level models (also known as mixed effects models or hierarchical models)

Figure 3 illustrates the set-up of a hypothetical longitudinal study designed to examine the impact of educational and workplace factors on physician capacity for compassion over time. The specific goal is to determine whether early educational influences wane over time and are successively replaced by the values of the current workplace culture or continue to provide a baseline for compassionate behaviour. The study assesses each physician's capacity for compassion at multiple time points during the study (e.g., annually during their education, internship, and residency, at the beginning of post-residency employment, and annually thereafter for 5 years), along with a battery of factors of interest.

As noted above, in a longitudinal study, repeated measurements on the same study participant, in this case individual physicians participating in the study, tend to be correlated. Furthermore, in this set-up, the compassion responses of physicians who studied at the same medical schools, trained at the same hospitals, or work at the same hospitals may be more similar than those of physicians from different environments, particularly if the former also share specialties or past mentors. If these dependencies are not accounted for when analyzing the data, confidence intervals around estimates of association may be misleadingly narrow. Multi-level models can account for these dependencies by explicitly partitioning the variation in compassion

responses across groups, allowing us to model both overall covariate effects (averaged across the study population) and the additional between-group variation that arises due to shared within-group environments, as well as the within-subject variation that arises due to repeated measurements. In our example, using multi-level models to analyze study data allows for examination of the influence of factors at multiple levels—e.g., the policy priorities of the administration at the city level, the degree of affluence of the population served at the hospital level, the degree of focus on compassionate practice at the level of the hospital department, and the socioeconomic background of the individual physician—on physician capacity for compassion, while accounting for clustering arising from shared medical schools, residency programs, workplace, and mentors.

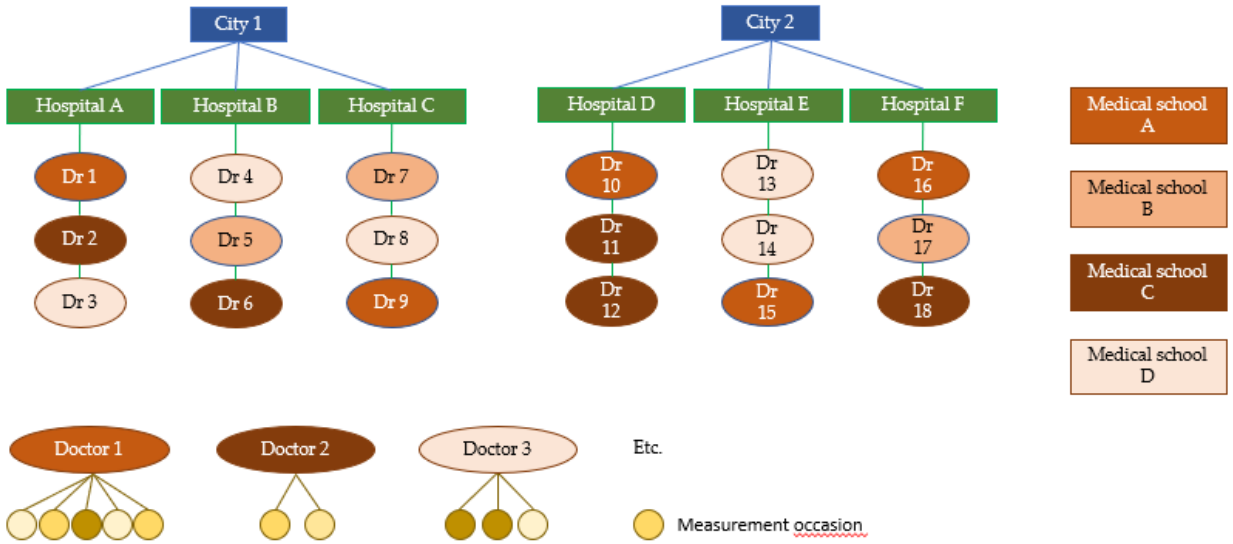

Figure 3. Illustration of a longitudinal study designed to assess educational, workplace, and regional factors that could influence physician capacity for compassion over time. For simplicity, only medical school (brown) and current city (blue) and hospital (green) of employment are included in the figure. Personal level of compassion measured on successive occasions is shown in yellow.

Although multi-level models can accommodate variability in the length of follow-up from one participant to the next (e.g., due to loss to follow-up or death), such variability does need to be accounted for in the analysis. In contrast, multi-level models are limited by their inability to accommodate time-varying covariates, particularly those that are not cumulative or irreversible (e.g., degree of religious or spiritual practice, which can fluctuate over time), nor can they handle situations in which covariates, including the exposure of interest, and outcomes evolve dynamically and may influence one another over time. As a result, multi-level models are best suited to studies with short timelines or independent variables that only vary infrequently or are fixed. A different method would be preferable if the goal of the study was to evaluate or account for the impact of reversible, short-term, or evolving influences on the capacity for compassion.

4.2 Methods that accommodate time-varying covariates

The handling of time-varying exposures and confounders has been an area of intense focus in the epidemiological literature, and the methods for causal inference that have been developed to address time-varying covariates can be used to advantage in compassion research. Consider a hypothetical study of the experience of adversity on the capacity for compassion, where social

network and neighbourhood variables have been deemed important potential confounders. In this design, all variables are assessed repeatedly over the study period, allowing researchers to study how adverse life events (e.g., their accumulation or the occurrence of multiple types) affect the evolution of capacity for compassion in the presence of potentially fluctuating levels of social power and varying neighbourhood influences.

As a first step to understanding the interplay of covariates of potential importance when studying an outcome of interest, a commonly used tool in epidemiologic research is the direct acyclic graph (Etminan et al., 2020), which displays assumptions about the associations between variables, with arrows to denote the assumed direction of the associations. Figure 4 presents such a directed acyclic graph for the hypothetical study of adverse life events, social network, and compassion. One example of the complex relationships uncovered by this graph is the fact that social network at 20 years of age is both a confounder of the association between an individual's capacity for compassion at age 25 years and their experience of adversity at that same age and also one the pathways through which childhood adversity influences capacity for compassion at age 25 years. Given these complex relationships, standard regression models that simply included the social network as an adjustment variable would yield biased results, because the adjustment would block an important causal pathway.

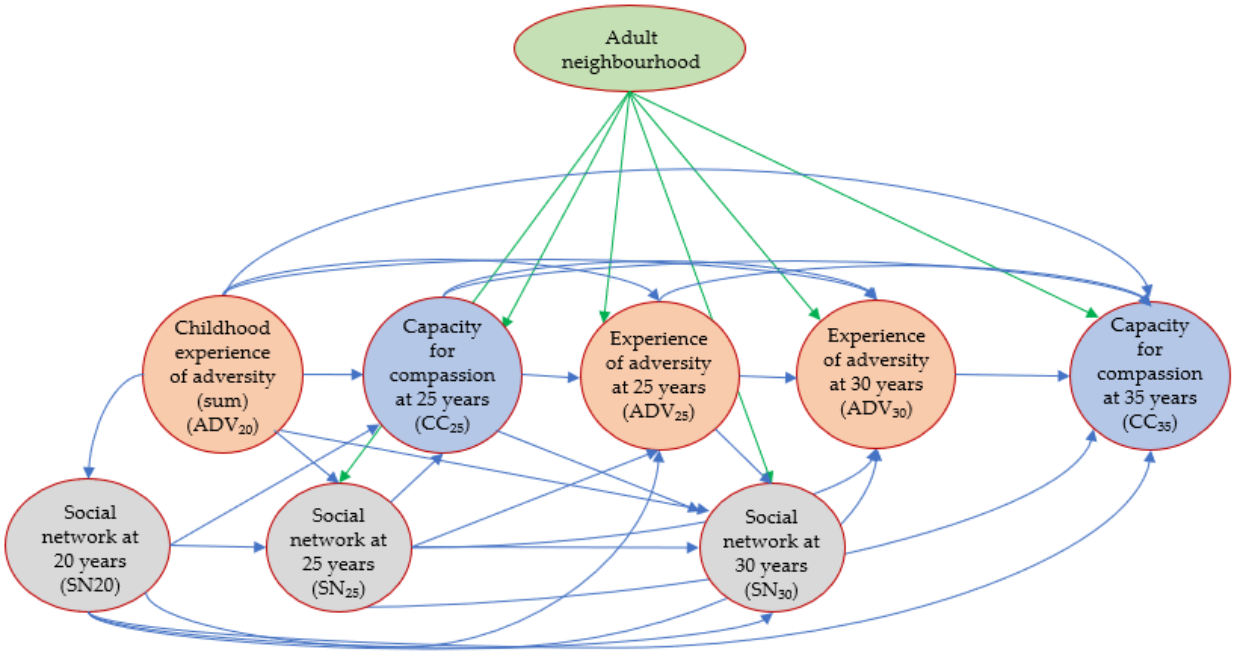

Figure 4. Simplified directed acyclic graph illustrating the hypothesized relationships between the main variables of interest in a hypothetical study of the experience of adversity in childhood and early adulthood (orange) on the capacity for compassion in the same periods (blue), where social network (grey) and adult neighbourhood (green) variables have been deemed important potential confounders. Capacity for compassion was assessed at age 25 years and again at age 35 years, and representations of both social network and cumulative experience of adversity were assessed at 20, 25, and 30 years

If the goal is to estimate short-term total effects, such as the effect of cumulative adversity by age 20 years on the capacity for compassion at 25 years of age, conventional methods such as sequential conditional mean models (Keough et al., 2018) should be sufficient to accommodate time-varying covariates. However, if the aim is to separate the effect of adversity that is mediated

by social network from the overall effect or to look at long-term effects, such as the effect of childhood adversity on capacity for compassion at 35 years of age, more advanced methods are needed. Many methods have been developed specifically to accommodate time-varying exposures and confounders, including marginal structural models, the g-computation formula, nested structural models, and targeted maximum likelihood estimation (Daniel et al., 2013; Newsome et al., 2018), the most widely-used of which are marginal structural models.

Compared with standard multi-level models, marginal structural models do not adjust for time-dependent confounders directly but use inverse probability weighting to create a pseudo-population where the exposure groups are balanced with respect to measured confounders (Robins et al., 2000). Marginal structural models provide flexibility in the handling of time-varying factors and repeated outcome measurements and can be modified to account for censoring due to loss to follow-up (VanderWeele et al., 2009; Wu et al., 2022). However, this class of models requires a high level of expertise to be used correctly and can be computationally intensive, since standard errors must be estimated by bootstrapping. In addition, marginal structural models (and the other types of models mentioned in the previous paragraph) are founded on strong assumptions that may not hold for a given study. In particular, it is assumed that there is no unmeasured confounding i.e., that all potential confounding variables have been identified, measured, and included in the model.

4.3 Machine learning methods

Machine learning—the use of high-performance computing to "learn" the most logical patterns and connections in a dataset in completely hypothesis-free fashion—holds promise as a way to reduce complex, high-dimensional data that vary over time to manageable clusters of factors that are associated with an outcome of interest. By pinpointing the most informative variables in a variable-rich dataset and/or summarizing related variables into a smaller number of meaningful dimensions, machine learning methods can identify previously unsuspected complex interactions among covariates, reveal axes of heterogeneity, accommodate time-varying fluctuations in factor levels, and produce useful knowledge from datasets with thousands of covariates, even if these are only measured on modest numbers of study participants. Different machine learning approaches are suited to different goals (Box 1). For instance, tree-based methods like random forests and gradient boosting can uncover complex interactions and highlight the most informative variables in large datasets (Breiman, 2001; Strobl et al., 2009). Clustering algorithms such as k-means and latent class models can reveal hidden outcome subgroups and their determinants (Jain, 2010; Naldi & Cazzaninga, 2020). Techniques for dimension reduction such as principal component analysis, factor models, and autoencoders can summarize many covariates into a smaller set of dimensions (Jolliffe & Cadima, 2016; Kingma & Welling, 2019). Extensions of clustering methods for time-varying covariates, recurrent neural networks, and dynamic Bayesian models can capture temporal patterns when covariate values vary over the study period.

Box 1. Suggested reading on the statistical methods highlighted as potentially useful in studies of the determinants of compassion

Multi-Level Models

Gelman, A., & Hill, J. (2006). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge: Cambridge University Press. [A comprehensive overview of multi-level models, with practical examples.]

Box 1. Suggested reading on the statistical methods highlighted as potentially useful in studies of the determinants of compassion (cont.)

Methods That Accommodate Time-Varying Covariates

Keogh, R.H., Daniel, R.M., VanderWeele, T.J., & Vansteelandt, S. (2018). Analysis of longitudinal studies with repeated outcome measures: Adjusting for time-dependent confounding using conventional methods. *American Journal of Epidemiology*. 187: 1085-92. [An overview of how standard regression methods can be adapted to provide unbiased estimation in the presence of time-varying confounding, yielding an estimation approach known as sequential conditional mean modelling.]

Nandi, A., Glymour, M.M., Ichiro, K., & VanderWeele, T. (2012). Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology*. 23: 223-32. [An example of the advantages of marginal structural models for investigating determinants of chronic disease when associations may be subject to time-varying confounding.]

Newsome, S.J., Keogh, R.H., & Daniel, R.M. (2018). Estimating long-term treatment effects in observational data: A comparison of the performance of difference methods under real-world uncertainty. *Statistics in Medicine*. 37: 2367-2390. [An overview of inverse probability weighting in marginal structural models, history-adjusted marginal structural models, sequential conditional mean models, the g-computation formula, and g-estimation of structural nested models, with discussion of their relative merits and limitations in the context of modelling treatment of cystic fibrosis.]

Robins, J., Hernán, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*. 11: 550-60. [The original introduction of marginal structural models and inverse-probability-of-treatment weighted estimators to deal with time-dependent confounders in observational studies.]

VanderWeele, T.J. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 20: 18-26. [Extension of marginal structural models to describe the extent to which the association between a given determinant and the outcome of interest is mediated by an intermediate variable.]

Wu, Y., Langworthy, B., & Wang, M. (2022). Marginal structural models for multilevel clustered data. *Biostatistics*. 23: 1056-73. [A statistical description of marginal structural models and some extensions that can be used to handle features commonly encountered in longitudinal studies.]

Machine Learning Methods

Random Forests

Breiman, L. (2001). Random forests. *Machine Learning*. 45: 5-32. [The seminal – and highly statistical – paper introducing random forests.]

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*. 14: 323-348. [An accessible explanation of how random forests and other recursive partitioning methods work.]

Box 1. Suggested reading on the statistical methods highlighted as potentially useful in studies of the determinants of compassion (cont.)

Support Vector Machines
Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*. 20: 273-297. [The original – and highly statistical – paper introducing support-vector machines.]

Noble, W. (2006). What is a support vector machine? *Nature Biotechnology*. 24: 1565-1567. [An accessible explanation of how support vector machines work, with a real-life example.]

Principal Component Analysis
Jolliffe, I.T., & Cadima, J. (2016). Principal component analysis: a review and recent developments *Philosophical Transactions. Series A: Mathematical, Physical and Engineering Sciences*. 374: 20150202 [A review of principal component analysis and its adaptations for a variety of data types. Statistical with real-world examples.]

Latent Class Models
Naldi, L., & Cazzaniga, S. (2020). Research techniques made simple: Latent class analysis. *Journal of Investigative Dermatology*. 140: 1676-1680. [A readable primer on latent class analysis, with examples from dermatology.]

K-Means Clustering
Jain, A.K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 31: 651-666. [An overview of cluster analysis and well-known clustering methods and a discussion of important considerations and challenges in designing clustering algorithms.]

Variational Autoencoders
Kingma, D.P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*. 12: 307-392. [The statistical principles behind variational autoencoders.]

These properties suggest that machine learning methods could be harnessed in early epidemiological studies of compassion to identify candidate determinants of the capacity for compassion for further study and generate hypotheses. The emerging machine learning landscape is varied, and so too are the strengths and limitations of different approaches, but challenges common to many methods include sensitivity to data quality (the problem of "garbage in, garbage out"), limited generalizability of results beyond the data used to train the models, and high computational demands. For these reasons, machine learning has thus far mainly been applied in the exploratory stages of research, while its applications in more rigorous causal studies remain limited. However, the emerging area of causal machine learning (Feuerriegel et al., 2024), which encompasses methods such as causal forests and targeted maximum likelihood estimation, is developing rapidly and robust approaches are likely to emerge in the next decade. The study of compassion is still in its early stages, and while causal machine learning approaches are not fully realized at present, they promise to provide useful methods for hypothesis testing and causal inference at later stages of compassion research. Similarly, recently developed methods for causal mediation analysis boosted by machine learning (methods for identifying and evaluating the assumptions under which mediators are established to play a causal role in an observed association, e.g., Farbmacher et al., 2022) may also help to untangle how various potential determinants interact to influence the capacity for compassion.

4.4 Other possibilities

Methods used to predict prognosis and treatment outcomes in chronic diseases with non-linear progression (e.g., multiple sclerosis, severe depression) may also be useful when analyzing datasets in which capacity for compassion is measured repeatedly. These methods, which include discrete event simulation, system dynamics modelling, and agent-based modelling, as well as machine learning and deep learning methods, have previously been used to model diseases with prognostic outlooks that at any given time can include remission, relapse, and progression with no set pattern or time scale (Piena et al., 2022; Reifsnider et al., 2022; Brice et al., 2023; Sabarre, 2024).

5. Alternative study designs potentially suitable for studying compassion

Variations on classical epidemiologic study designs might be particularly well-suited to addressing issues that complicate the study of the determinants of compassion. In this section, we highlight two study designs used infrequently in compassion studies to date.

5.1 Sibling comparison design

A variety of childhood factors and experiences are thought to influence the capacity for compassion throughout the lifecourse (Addiss et al., 2022). Under normal circumstances, epidemiologic studies of early-life determinants of compassion must therefore identify, measure, and mutually adjust for all such factors, no easy task when studying outcomes years downstream. In this situation, the *sibling comparison design*, a matched cohort study variant that compares siblings discordant for an exposure of interest, might prove useful (Frisell et al., 2012; Keyes & Susser, 2023). The capacity for compassion could be compared in pairs of full siblings discordant for e.g., completion of a university degree, experience of financial or housing insecurity as a young adult, or residence after parental divorce. Because full siblings are assumed to have been brought up in the same household(s) by the same parents and exposed to most of the same influences, values, and childhood events, the study design minimizes the potential confounding influence of shared genetics and early-life experiences and influences, whether measured or unmeasured. This design feature renders the sibling comparison study most suitable for examining the role of factors in childhood, adolescence, and young adulthood on the outcome of interest, as siblings' influences presumably begin to diverge once they leave home, increasing the potential for unshared confounders that would need to be identified, measured and accounted for in the analysis. The sibling comparison design assumes that the outcome, in this case the capacity for compassion, is not hard-wired (i.e., genetically determined and immutable), that there are no unmeasured non-shared confounders, and that the "exposed" sibling's experience of the independent variable of interest does not have an influence on the "unexposed" sibling's capacity for compassion (i.e., that there is no spillover influence of the factor of interest from one sibling to the other). As with analyses of data from other types of matched studies, it is necessary to account for the within-family matching when analysing data from sibling comparison studies. Consequently, depending on how the outcome is defined, the most common methods used to analyze data from sibling comparison studies include fixed-effects ordinary least squares modelling, conditional logistic regression, and stratified Cox regression.

5.2 Difference-in-difference design

As previously noted, in some instances, randomized experimental studies of public health interventions are not practical or even possible. Testing the impact of an intervention to increase

compassion on a large scale—e.g., studying the effects of workplace-, community-, city-, or even nationwide policies on the capacity for compassion at a population level—might be one such instance. Despite their limitations, quasi-experimental designs, the *difference-in-difference design* in particular, could serve as useful alternative study designs for investigating the impact of group- or population-level compassion interventions (Wing et al., 2018; Caniglia & Murray, 2020; Capili & Anastasi, 2024).

In the difference-in-difference design, a "treatment" group or population is selected to receive the intervention while another group with similar characteristics is selected to be the comparison group. Both groups complete a pre-intervention assessment of the outcome measure(s) of interest. The treatment group then receives the intervention, while the comparison group remains untreated. Finally, the outcome measures are reassessed in both groups. By comparing outcome data from the two groups before and after the intervention, the design allows for the estimation of the average effect of treatment among those who received the intervention.

For example, the design could be used to compare the change in the capacity for compassion over time at workplaces A and B, where workplace A's leadership implemented a partnership with the surrounding community to provide mentorship to local youth, while workplace B did not. By analyzing data from compassion assessments in both workplaces at the beginning and end of the study period (i.e., before workplace A initiated its program and a reasonable amount of time after the program was well-established, respectively), this design would allow researchers to estimate the average effect that implementation of the community mentorship program had on the capacity for compassion in employees at workplace A. In essence, the method compares the average change in capacity for compassion among workplace A's employees during the study period had they all been involved in the mentorship program with the average change in compassion in the same group had none of them taken part in the mentorship initiative.

However, in interpreting the results of a difference-in-difference study, it must be noted that certain assumptions must be met to yield an unbiased estimate of the effect of implementing the intervention and to suggest that any observed changes in outcome were due to the intervention. First, the method assumes that the population-level exposure is specific and well-defined. The design is well-suited to studying the impact of the specific program implemented at workplace A, compared with the absence of a mentorship program; it would not be appropriate for a study comparing the implementation of different programs at different workplaces. Second, the method assumes that any unmeasured determinants of the outcome (here, capacity for compassion) do not vary over the study period or do not vary between workplaces; in other words, this assumption of "parallel trends" requires that in the absence of the intervention (the implementation of the mentorship program at workplace A), the difference in compassion levels at workplaces A and B would have remained constant over the study period. Third, the study design assumes that workplace A's decision to implement the mentorship program, and workplace B's decision (active or passive) not to take similar action, were not related to the capacity for compassion at workplace A and B, respectively, at the start of the study period. A final assumption is that all persons in the target population being studied were actually eligible to receive all levels of the exposure—that is, that there was no external requirement for workplace A but not workplace B to implement its policy change and no barrier to workplace B implementing a similar policy.

Data from difference-in-difference studies are typically analyzed with standard regression models that include an interaction term between time and exposure group (the difference-in-difference estimate). Robust standard error estimators should be used to account for within-group clustering.

Box 2. Suggested reading on the study designs highlighted as potentially useful in studies of the determinants of compassion

Longitudinal Cohort Studies

Guralnik, J.M., & Kritchevsky, S. (2011). Translating research to promote healthy aging: the complimentary role of longitudinal studies and clinical trials. *Journal of the American Geriatric Society*. 58: S337-42. [A clear presentation of the advantages, disadvantages, and complementarity of longitudinal cohort studies (observational) and randomized clinical trials (experimental), with aging studies used as an example.]

Zouhair, J. (2023). Longitudinal prospective cohort studies: a research method primer. *Undergraduate Research in Natural and Clinical Science and Technology Journal*. 7: 1-5. [An overview of the key features of longitudinal cohort studies.]

Sibling Comparison Design

Frisell, T., Öberg, S., Kuja-Halkola, R., & Sjölander, A. (2012). Sibling comparison designs. Bias from non-shared confounders and measurement error. *Epidemiology*. 23: 713-20. [A statistical discussion of the limitations of the sibling comparison design, how to mitigate potential bias from non-shared confounders, and some non-statistical rules of thumb for when the design could be useful.]

Keyes, K.M., & Susser, E. (2023). Uses and misuses of sibling designs. *International Journal of Epidemiology*. 52: 336-341. [An accessible, non-statistical discussion of when it is appropriate to use the sibling comparison design and when it would be better to use another method.]

Difference-in-Difference Design

Caniglia, E.C., & Murray, E.J. (2020). Difference-in-difference in the time of cholera: A gentle introduction for epidemiologists. *Current Epidemiology Reports*. 7: 203-211. [An overview of the difference-in-difference method, with an example from classic epidemiology.]

Capili, B., & Anastasi, J.K. (2024). An introduction to the quasi-experimental design (nonrandomized design). *American Journal of Nursing*. 124: 50-52. [A gentle introduction to three quasi-experimental designs.]

Wing, C., Simon, K., & Bello-Gomez, R.A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*. 39: 453-469. [An in-depth statistical discussion of the difference-in-difference method and its extensions to more complicated cases than the simple study outlined in our example.]

6. Challenges in using methods from chronic disease epidemiology to study compassion

Despite our conviction that methods from chronic disease epidemiology are well-suited to epidemiological studies of compassion in theory, in practice several challenges need to be addressed to allow these methods to be applied to real-life study data.

6.1 Defining compassion

Epidemiology is a data-based science that quantifies the contributions of study factors to absolute and relative risks of the outcome of interest. Successful epidemiological studies of compassion will therefore hinge on our ability to define and quantify compassion. Recent work has frequently used compassion scores generated by summarizing study participants' responses to questionnaire items at one or more points in time. For example, in a study of potential genetic

determinants of dispositional compassion, Dobewall and colleagues administered the Temperament and Character Inventory, a personality inventory consisting of 240 items to which participants respond using Likert scales or true/false answers (Cloninger et al., 1993); the authors then generated compassion scores based on responses to the Compassion subscale within the inventory's Cooperativeness dimension (Dobewall et al., 2021). However, more sophisticated measures of compassion will be required to allow realistic modelling of its determinants. Furthermore, how we define compassion reflects how we believe compassion "works", which in turn informs our choice of study design and statistical methods. Elsewhere in this issue, Ozawa-de Silva & Mascaro outline what is needed to produce a framework for understanding and measuring compassion in ways suitable for use in epidemiological studies (Ozawa-de Silva & Mascaro, 2026).

6.2 Time scale and availability of sufficiently granular data

As noted in previous sections, the determinants of compassion most likely range from short, episodic effects to stable, long-term influences and include factors that are fixed, those that vary transiently, and those that vary in irreversible ways. The aspects of compassion to be investigated will therefore dictate a study's time scale and the data granularity required with respect to time; they may also limit the usefulness of existing datasets.

6.3 Compassion as a transmissible trait

While compassion shares many similarities with chronic diseases, it also has features reminiscent of infectious diseases, including the potential for transmission (spread) through mechanisms such as formal education, informal role modeling, social nudging, adoption of workplace or community values and norms, and social media influences. In addition, the capacity for compassion may mimic the capacity for infectious disease, with individuals cycling through states of susceptibility (when outside influences may exert an effect on capacity for compassion), active "infection" (when the capacity for compassion is heightened or lowered relative to the initial state), "immunity" (when the change in capacity for compassion is sustained and cannot be altered), and reversion to the susceptible state (return to an earlier capacity for compassion once the outside influence has faded). Integrating transmissible aspects of compassion into epidemiologic models will require forays into methods used to model the transmission of infectious diseases and other methods for network analysis, further supporting the potential of methods from the artificial intelligence world to help in explaining the determinants of compassion.

7. Conclusion

By using epidemiologic study designs and analytic methods to identify the determinants of compassion and quantify their impact and interactions, we will pave the way for rigorous, quantitative assessment of potential interventions to increase compassion in populations. If challenges in defining the capacity for compassion can be overcome, methods typically used in chronic disease epidemiology, particularly methods that can accommodate time-varying covariates and emerging machine learning methods, promise to make important contributions to the study of compassion and its determinants.

Authors

Heather A. Boyd
Department of Congenital Diseases, Statens Serum Institut, Copenhagen, Denmark
https://orcid.org/0000-0001-6849-9985
hoy@ssi.dk

Anna Damkjær Laksafoss
Department of Data Science and Artificial Intelligence in Health, Statens Serum Institut, Copenhagen, Denmark
https://orcid.org/0000-0002-9898-2924

Henrik Hjalgrim
Danish Cancer Institute, Copenhagen, Denmark; Department of Data Science and Artificial Intelligence in Health, Statens Serum Institut, Copenhagen, Denmark; Department of Haematology, Rigshospitalet, Copenhagen, Denmark; Department of Clinical Medicine, Copenhagen University, Copenhagen, Denmark
https://orcid.org/0000-0002-4436-6798

Gry Juul Poulsen
Center for Molecular Prediction of Inflammatory Bowel Disease (PREDICT), Department of Clinical Medicine, Aalborg University Copenhagen, Denmark
https://orcid.org/0000-0003-2744-2469

Author contribution statement

Boyd pitched the original idea for the article, which the other three co-authors then helped to refine. Boyd wrote the first draft of the article, with input from Laksafoss and Poulsen. All co-authors provided substantive edits and comments, which Boyd then incorporated into the final version of the article.

Funding

While writing this article, Boyd was supported by the Novo Nordisk Foundation (grant no. NNF19OC0054286) and Poulsen was supported by a grant from the Danish National Research Foundation (grant no. DNRF148).

Conflict of interest statement

The authors declare that they have no conflict of interest.

AI statement

AI was not used in drafting this article.

Publishing Timeline

Received 31 August 2025
Revised version received 1 December 2025
Accepted 3 December 2025
Published 4 March 2026

References

Addiss, D.G., Richards, A., Adiabu, S., Horwath, E., Leruth, S., Graham, A.L., & Buesseler, H. (2022). Epidemiology of compassion: A literature review. *Frontiers in Psychology*. 13: 992705. https://doi.org/10.3389/fpsyg.2022.992705

Breiman, L. (2001). Random forests. *Machine Learning.* 45: 5-32. https://doi.org/10.1023/A:1010933404324

Brice, S.N., Harper, P.R., Garner, D., & Behrens, D.A. (2023). Modeling disease progression and treatment pathways for depression jointly using agent based modeling and system dynamics. *Frontiers in Public Health*. 10: 1011104. https://doi.org/10.3389/fpubh.2022.1011104

Brown, J.C., Gerhardt, T.C., & Kwon, E. (Updated 2023 Jan 23). Risk factors for coronary artery disease. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. Available from: Risk Factors for Coronary Artery Disease - StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK554410 Accessed August 2025.

Butts, M.M., Lunt, D.C., Freling, T.L., & Gabriel, A.S. (2019). Helping one or helping many? A theoretical integration and meta-analytic review of the compassion fade literature. *Organizational Behavior and Human Decision Processes*. 151: 16-33. https://doi.org/10.1016/j.obhdp.2018.12.006

Caniglia, E.C., & Murray, E.J. (2020). Difference-in-difference in the time of cholera: A gentle introduction for epidemiologists. *Current Epidemiology Reports*. 7: 203-211. https://doi.org/10.1007/s40471-020-00245-2.

Capili, B., & Anastasi, J.K. (2024). An introduction to the quasi-experimental design (nonrandomized design). *American Journal of Nursing*. 124: 50-52. https://doi.org/10.1097/01.NAJ.0001081740.74815.20

Cloninger, C.R., Svrakic, D.M., & Przybeck, T.R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry*. 50: 975-990. https://doi.org/10.1001/archpsyc.1993.01820240059008

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*. 20: 273-297. https://doi.org/10.1007/BF00994018

Daniel, R.M., Cousens, S.N., De Stavola, B.L., Kenward, M.G., & Sterne, J.A.C. (2013). Methods for dealing with time-dependent confounding. *Statistics in Medicine*. 32: 1584-1618. https://doi.org/10.1002/sim.5686

Dobewall, H., Saarinen, A., Lyytikäinen, L.P., Keltikangas-Järvinen, L., Lehtimäki, T., & Hintsanen, M. (2021). Functional polymorphisms in oxytocin and dopamine pathway genes and the development of dispositional compassion over time: The Young Finns Study. *Frontiers in Psychology*. 12: 576346. https://doi.org/10.3389/fpsyg.2021.576346

Etminan, M., Collins, G.S., & Mansournia, M.A. (2020). Using causal diagrams to improve design and interpretation of medical research. *Chest*. 158, Supplement 1: S21-S28. https://doi.org/10.1016/j.chest.2020.03.011

Farbmacher, H., Huber, M., Lafférs, L., Langen, H., & Spindler, M. (2022). Causal mediation analysis with double machine learning. *The Econometrics Journal*. 25: 277-300. https://doi.org/10.1093/ectj/utac003

Feuerriegel, S., Frauen, D., Melnychuk, V., Schweisthal, J., Hess, K., Curth, A., Bauer, S., Kilbertus, N., Kohane, I.S., & van der Schaar, M. (2024). Causal machine learning for predicting treatment outcomes. *Nature Medicine*. 30: 958-968. https://doi.org/10.1038/s41591-024-02902-1

Frisell, T., Öberg, S., Kuja-Halkola, R., & Sjölander, A. (2012). Sibling comparison designs. Bias from non-shared confounders and measurement error. *Epidemiology*. 23: 713-20. https://doi.org/10.1097/EDE.0b013e31825fa230

Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

Gu, J., Baer, R., Cavanagh, K., Kuyken, W., & Strauss, C. (2020). Development and psychometric properties of the Sussex-Oxford compassion scales (SOCS). *Assessment.* 27: 3-20. https://doi.org/10.1177/1073191119860911

Guralnik, J.M., & Kritchevsky, S. (2011). Translating research to promote healthy aging: the complimentary role of longitudinal studies and clinical trials. *Journal of the American Geriatric Society*. 58: S337-42. https://doi.org/10.1111/j.1532-5415.2010.02938.x

Jain, A.K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 31: 651-666. https://doi.org/10.1016/patrec.2009.09.011

Jha, A., Minocha, R., Patel, H., Akhouri, S.S., Mishra, T., Barik, S., Firoz, S., Singh, M.K., Kashyap, A., Das, S., Sethi, A., Rajan, T., Gautam, D., Singh, M.K., Chakraborty, K., Reddy, N.S., Das, A., Varma, K., &

Mahapatra, T. (2026). Fostering compassionate public healthcare leadership towards sustainability: Evidence from Bihar, India. *International Journal of Wellbeing*. 16(2), 5605, 1-28. https://doi.org/10.5502/ijw.v16i2.5605

Jolliffe, I.T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions. Series A: Mathematical, Physical and Engineering Sciences*. 374: 20150202 https://doi.org/10.1098/rsta.2015.0202

Keogh, R.H., Daniel, R.M., VanderWheele, T.J., & Vansteelandt, S. (2018). Analysis of longitudinal studies with repeated outcome measures: Adjusting for time-dependent confounding using conventional methods. *American Journal of Epidemiology*. 187: 1085-1092. https://doi.org/10.1093/aje/kwx311

Keyes, K.M., & Susser, E. (2023). Uses and misuses of sibling designs. *International Journal of Epidemiology*. 52: 336-341. https://doi.org/10.1093/ije/dyac187

Kingma, D.P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*. 12: 307-392. https://doi.org/10.1561/2200000056

Li, H., Zheng, C., Zhang, Y., Yang, H., & Li, J. (2023). The directed acyclic graph helped identify confounders in the association between coronary heart disease and pesticide exposure among greenhouse vegetable farmers. *Medicine (Baltimore)*. 102: e35073. https://doi.org/10.1097/MD.0000000000035073

Lim, D., & DeSteno, D. (2020). Past adversity protects against the numeracy bias in compassion. *Emotion.* 20: 1344-1356. https://doi.org/10.1037/emo0000655

Lovette-Colyer, M. (2013). Cultivating compassion in undergraduate college students: Rhetoric or reality? Available at: https://digital.sandiego.edu/dissertations/846. Accessed August 2025. https://doi.org/10.22371/05.2013.010

Naldi, L., & Cazzaniga, S. (2020). Research techniques made simple: Latent class analysis. *Journal of Investigative Dermatology*. 140: 1676-1680. https://doi.org/10.1016/j.jid.2020.05.079

Nandi, A., Glymour, M.M., Ichiro, K., & VanderWeele, T. (2012). Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology*. 23: 223-32. https://doi.org/10.1097/EDE.0b013e31824570bd

Newsome, S.J., Keogh, R.H., & Daniel, R.M. (2018). Estimating long-term treatment effects in observational data: A comparison of the performance of different methods under real-world uncertainty. *Statistics in Medicine*. 37: 2367-2390. https://doi.org/10.1002/sim.7664

Noble, W. What is a support vector machine? *Nature Biotechnology*. 24: 1565-1567. https://doi.org/10.1038/nbt1206-1565

Ozawa-de Silva, B.R., & Mascaro, J.S. (2026). What's next for compassion science? A multi-scalar framework for measurement and public health action. *International Journal of Wellbeing*. 16(2), 5635, 1-24. https://doi.org/10.5502/ijw.v16i2.5635

Piena, M.A., Kroep, S., Simons, C., Fenwick, E., Harty, G.T., Wong, S.L., & van Hout, B.A. (2022). An innovative approach to modelling the optimal treatment sequences for patients with relapsing-remitting multiple sclerosis: Implementation, validation, and impact of the decision-making approach. *Advances in Therapy*. 39: 892-908. https://doi.org/10.1007/s12325-021-01975

Reifsnider, O.S., Kansal, A.R., Wanner, C., Pfarr, E., Koitka-Weber, A., Brand, S.B., Stargardter, M., Wang, C., Kuti, E., & Ustyugova, A. (2022). Cost-effectiveness of empagliflozin in patients with diabetic kidney disease in the United States: Findings based on the EMPA-REG OUTCOME trial. *American Journal of Kidney Disease*. 79: 796-806. https://doi.org/10.1053/j.ajkd.2021.09.014

Robins, J., Hernán, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*. 11: 550-60. https://doi.org/10.1097/00001648-200009000-00011

Sabarre, J. (2024). Advancing predictive modeling of multiple sclerosis progression through big data analytics and machine learning. *Oruen – The CNS Journal*. 10: 72-81. Available at: https://www.oruen.com/wp-content/uploads/2024/12/72-81_Oruen-Vol-10-Issue-2_Digital.pdf Accessed 31 August 2025.

Salzberg, S. (1995). The nature of compassion. Available at: https://www.buddhistinquiry.org/article/the-nature-of-compassion/ Accessed 28 August 2025.

Strauss, C., Taylor, B.L., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical Psychology Review.* 47: 15-27. https://doi.org/10.1016/j.cpr.2016.05.004

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*. 14: 323-348. https://doi.org/10.1037/a0016973.

VanderWeele, T.J. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 20: 18-26. https://doi.org/10.1097/EDE.0b013e31818f69ce

van Kleef, G.A., Oveis, C., van der Löwe, I., LuoKogan, A., Goetz, J., & Keltner, D. (2008). Power, distress, and compassion: turning a blind eye to the suffering of others. *Psychological Science.* 19: 1315-1322. https://doi.org/10.1111/j.1467-9280. 2008.02241.x

Wing, C., Simon, K., & Bello-Gomez, R.A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*. 39: 453-469. https://doi.org/10.1146/annurev-publhealth-040617-013507

Wu, Y., Langworthy, B., & Wang, M. (2022). Marginal structural models for multilevel clustered data. *Biostatistics*. 23: 1056-73. https://doi.org/10.1093/biostatistics/kxac027

Zouhair, J. (2023). Longitudinal prospective cohort studies: a research method primer. *Undergraduate Research in Natural and Clinical Science and Technology Journal*. 7: 1-5. https://doi.org/10.26685/urncst.399