PNAS Nexus, 2024, 3, 1–16
https://doi.org/10.1093/pnasnexus/pgae050
Advance access publication 2 February 2024
Research Report

Forecasting drug-overdose mortality by age in the United States at the national and county levels

Lucas Böttcher [a,*], Tom Chou [b] and Maria R. D'Orsogna [b,c]

[a]Department of Computational Science and Philosophy, Frankfurt School of Finance and Management, 60322 Frankfurt am Main, Germany
[b]Department of Computational Medicine, University of California, Los Angeles, Los Angeles, CA 90095-1766, USA
[c]Department of Mathematics, California State University at Northridge, Los Angeles, CA 91330-8313, USA
*To whom correspondence should be addressed: Email: l.boettcher@fs.de
Edited By: Sandro Galea

Abstract

The drug-overdose crisis in the United States continues to intensify. Fatalities have increased 5-fold since 1999 reaching a record high of 108,000 deaths in 2021. The epidemic has unfolded through distinct waves of different drug types, uniquely impacting various age, gender, race, and ethnic groups in specific geographical areas. One major challenge in designing interventions and efficiently delivering treatment is forecasting age-specific overdose patterns at the local level. To address this need, we develop a forecasting method that assimilates observational data obtained from the CDC WONDER database with an age-structured model of addiction and overdose mortality. We apply our method nationwide and to three select areas: Los Angeles County, Cook County, and the five boroughs of New York City, providing forecasts of drug-overdose mortality and estimates of relevant epidemiological quantities, such as mortality and age-specific addiction rates.

Significance Statement

The drug-overdose epidemic in the United States continues to escalate, with fatalities increasing 5-fold since 1999 and reaching a record high of 108,000 individuals in 2021. The crisis is characterized by distinct waves of drug types, disproportionately affecting various demographic groups in specific geographical regions. One major component of designing effective interventions is forecasting age-specific overdose patterns in order to facilitate targeted prevention and preparedness efforts. To this end, we propose a forecasting approach that integrates observational data with an age-structured model of addiction and overdose mortality. Applying this method nationwide and in areas that are highly impacted by the overdose crisis, we provide robust drug-overdose mortality forecasts offering vital insights for effective interventions.

Introduction

The United States is currently experiencing one of its worst drug crises, with alarming increases in fatal overdose rates. According to data from the Center for Disease Control (CDC), over 108,000 persons died from drug overdose in 2021, the highest number ever recorded in a single year and a 17 % increase over the previous record high of 2020 (1). Most recent overdose deaths involve synthetic opioids such as fentanyl, psychostimulants such as methamphetamines and, to a lesser degree, prescription opioids such as oxycodone, and heroin (2). Many factors may have contributed to this surge, including the overall increased supply of synthetic, low-cost drugs (3, 4), the ease with which illegal substances may be purchased online (5–7), the uncontrolled mixing of drugs of different potency (8, 9), and societal changes leading to "deaths of despair" (10, 11). Although these elements have fueled high-risk drug use for quite some time, most of them have been exacerbated by the COVID-19 pandemic (12, 13).

Both the CDC and the National Center for Health Statistics have been systematically collecting information on overdose mortality since 1999, and, according to slightly different classifications, since 1979. The relevant data are publicly accessible through the CDC Wide-ranging Online Data for Epidemiologic Research (WONDER) portal which is updated at the end of each calendar year with final data associated to the prior year, resulting in a 1-year lag. Many groups have dissected these data by stratifying overdoses according to drug type, year, age, gender, race, and geography. These studies have revealed several spatiotemporal "overdose waves" across the United States, the emergence of new trends, demographic and geographical shifts, and social disparities (13–16).

While providing up-to-date snapshots and following the course of past overdose deaths helps shed light on the evolution of the drug epidemic (17), forecasting future overdose patterns, even in the short term, would allow for targeted preventive interventions and ensure the preparedness of public health agencies (14, 18, 19). Due to demographic, political and legislative heterogeneities across the United States, predictions on the national scale would be much less effective than those made at the more local level

Competing Interest: The authors declare no competing interest.
Received: September 14, 2023. **Accepted:** January 25, 2024

(20). Analysis at a more "granular" scale allows one to retain specific drug-market, socioeconomic, cultural, and geo-historical attributes that distinctly affect the drug-overdose trajectories. By not lumping these factors together, more realistic forecasts and tailored interventions (21, 22) can be developed. For instance, while drug overdoses may be decreasing at the regional level, certain counties, urban centers, or even zip codes within the same region may be experiencing surges among given subpopulations due to the introduction of new drugs to a circumscribed market.

As data collecting and manipulation capabilities have expanded, predicting overdose mortality (at any scale), while still in its infancy, has become a rapidly growing field. Given the many aspects of the drug addiction crisis, current studies rely on a variety of information including data on past overdoses, hospitalization, arrest, internet searches, painkiller prescription, and drug-seizure rates. Quantitative tools used in these endeavors include statistical regression, geospatial analyses, mathematical modeling, and machine learning (23–30).

In this paper, we advance the state of the art in drug-overdose forecasting by combining a mechanistic model describing age-stratified drug-overdose fatalities with recorded mortalities using data-assimilation techniques (31–33). The latter were first developed within the geological and atmospheric sciences to merge high-dimensional dynamical systems with large datasets to produce weather and climate forecasts. After decades of continuous improvement to both algorithms and computing infrastructure, modern operational weather forecasting centers are able to process about 10^7 observations per day (34). In addition to applications in climate dynamics, data assimilation has been used to estimate parameters in systems biology (35), to provide risk-dependent individual contact interventions during outbreaks (36), to identify patients with antibiotic-resistant bacteria in hospital wards (37), and to quantify the proportion of undocumented COVID-19 cases (38). One reason for the successful integration of mechanistic models with data-assimilation methods across different fields is that the algorithms are computationally efficient and provide good forecasts even when training data are sparse (39). Furthermore, since they are coupled to mechanistic models, data-assimilation methods allow one to estimate parameters that carry a physical or biological meaning and to follow their evolution over time. This interpretability, both of the parameters and of their dynamics, is very valuable for decision-making and formulating intervention policies. Finally, contrary to other techniques, data-assimilation methods produce interval estimates and not just point estimates. One can thus quantify confidence intervals (CIs) and accurately assess uncertainties and risks.

The mechanistic model that we use in this work is based on Kermack–McKendrick theory (40–44) and describes an age-structured, drug-using population. This group includes those who suffer from substance use disorder (SUD) as well as occasional or first-time users. Our model includes population aging, the age-dependent initiation of high-risk substance use, and drug-induced mortality. Using data assimilation to combine our drug-overdose model with data from CDC WONDER, we develop a forecasting tool for age-stratified drug-overdose mortality in the United States. In the next section, we illustrate the basic principles of our method by generating nationwide drug-overdose mortality forecasts and by extracting the time evolution of epidemiological quantities such as rates of initiation of drug use and mortality. We compare our predictions with overdose data for select past years and offer short-term projections for drug-overdose mortality. Similarly, we generate drug-overdose forecasts for select counties or metropolitan areas that display a large number of overdose fatalities: Los Angeles County, CA; Cook County, IL; and the five boroughs of New York City. Our forecasts show that age-structured population models combined with data-assimilation methods can produce reliable predictions of drug-overdose deaths both at the national and county levels. Our approach and its results can inform early warning systems, help tailor interventions, and prioritize resources distribution to areas most impacted by the current drug epidemic.

Results

Forecasting overdose fatalities in the United States

The Kermack–McKendrick model (40–43) is commonly used in mathematical epidemiology to describe the evolution of an age-structured population with age-dependent infection and recovery rates. Related structured population models have found utility in describing cell populations (45), demographics and birth control policies (46), the progression of infectious diseases (44) such as measles (47), tuberculosis (48), HIV (49), and COVID-19 (50) and more recently in the social sciences (51) and in studies of drug addiction (52–57). In this work, we combine an age-stratified model of overdose fatalities with corresponding observational data from the CDC WONDER database, using an ensemble Kalman filter (EnKF) (58) as data assimilation method (see Materials and methods for further information on the age-structured model, EnKF, and overdose data).

We model both the evolution of the population of those who use high-risk drugs and the number of fatal drug overdoses across different age classes in yearly increments. To estimate age-specific influx rates of high-risk drug consumers and age-specific mortality rates, we use the age-structured population data and fatal overdose data tallied by the CDC WONDER database as inputs to our EnKF. Since we consider the entire population of those who use high-risk drugs, overdose deaths may arise among those who suffer from SUD as well as intermittent or first-time users who accidentally overdose on contaminated doses or due to inexperience (59, 60). Figure 1a,b shows the evolution of the US population and overdose fatalities from 1999 (light blue) to 2021 (dark blue). Within this timeframe, the population between 0 and 85 years rose from 275 to 326 million individuals. Figure 1a shows that the largest increases occurred between 20 and 40 years and that the age-structured population distribution is marked by two characteristic peaks: one arising between 20 and 30 years and the other between 40 and 60 years. The age-structured fatal overdose distribution in Fig. 1b reveals that between 1999 and 2014, the largest proportion of overdose deaths occurred between 40 and 50 years. During a second phase, spanning from 2015 to 2019, overdoses peaked within the 35–40 year age group. A sudden surge in overdose fatalities is observed beginning in 2020; the onset of this third phase is concurrent with the advent of COVID-19. These three phases do not define rigid classifications; rather, they provide reference points to facilitate data interpretation and guide our analysis.

We use these qualitative observations to guide the development of our age-structured Kermack–McKendrick model. First, to allow for possible population shifts or shifts in the onset of high-risk substance use, we include two age-stratified influx rates in the shape of gamma distributions peaked at ages $a_1^{\max}$ and $a_2^{\max}$ with amplitudes r_1 and r_2, respectively. The two distributions may represent, for example, initiation of high-risk substance use among young adults and among middle-aged persons seeking relief from pain through prescription opioids. Large values of r_1

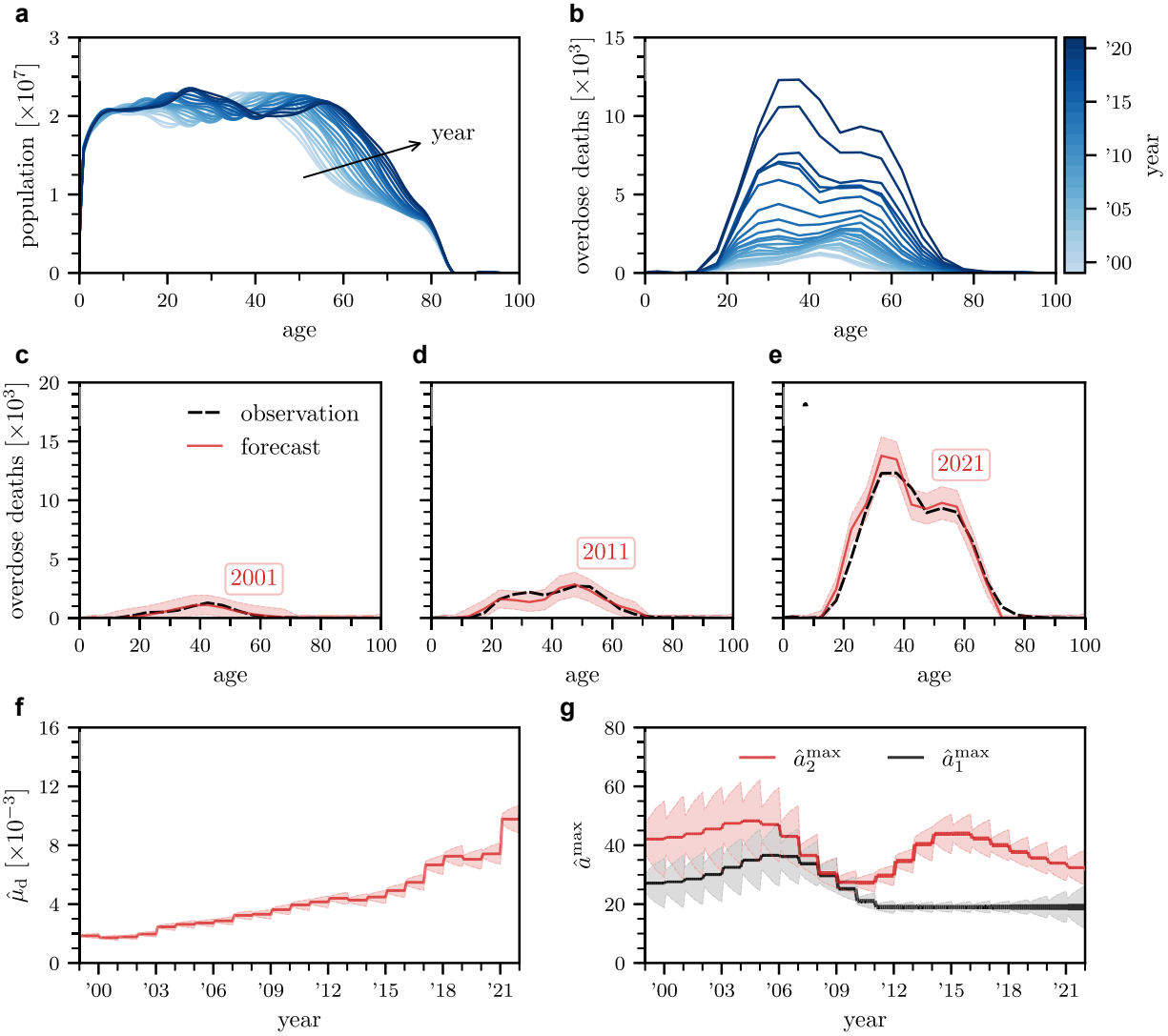

Fig. 1. Forecasting nationwide overdose fatalities. a and b) United States population and overdose deaths as a function of age (0–100 years) and time (1999–2021). c–e) Forecasts of overdose deaths as a function of age. Solid red curves and red-shaded areas indicate mean predictions and 3σ (i.e. 3-standard deviation) intervals, respectively. Observed fatalities are indicated by dashed black curves. Ages are binned in 5-year intervals and prediction values are displayed at the center of each bin interval. f) Evolution of the estimated drug-caused mortality rate $\hat{\mu}_d$ and the corresponding 3σ intervals. g) Evolution of the estimated ages for which the onset of high-risk drug use is largest and corresponding 3σ intervals. To account for potential population shifts, we utilize two age-stratified influxes peaked at $\hat{a}_1^{max}$, $\hat{a}_2^{max}$. The thickness of the curves is proportional to their respective magnitudes $\hat{r}_1$ and $\hat{r}_2$. As can be visually inferred, the influx peaked at the younger age $\hat{a}_1^{max}$ begins to carry more weight than its counterpart peaked at $\hat{a}_2^{max}$ around 2015, indicating a shift toward a preponderance of younger persons using high-risk drugs. We set a lower limit of 19 years for $\hat{a}_1^{max}$ to prevent unrealistically low ages of drug-use initiation. Filter updates occur in the beginning of each year.

compared to r_2 imply that the influx of new users occurs mostly through the distribution that is peaked at a_1^{max} and vice versa. Second, to take into account nonoverdose deaths among those who use high-risk drugs, we write the mortality rate of this population as an age-stratified baseline given by the Gompertz–Makeham–Siler approximation (61–65) to which an excess drug-induced mortality μ_d is added.

As typical in data assimilation, at each forecasting step, we use new overdose fatality data to update the system state and estimates of model quantities (such as the drug-induced mortality rate $\hat{\mu}_d$, the ages at which the influx rates of the population that uses high-risk drugs are maximal, $\hat{a}_1^{max}$, $\hat{a}_2^{max}$, and the amplitudes of the influx distributions $\hat{r}_1$, $\hat{r}_2$) and use these values for subsequent forecasts. Since final CDC WONDER data are available from 1999 to 2021, a forecast for drug-overdose fatalities in year Y is based on assimilated observational data between 1999 and Y – 1. Figure 1c–e displays age-stratified overdose forecasts (solid red curves) and corresponding observational data (dashed black curves) for the years 2001, 2011, and 2021. These are representative years selected from the three phases outlined above. In all panels, red-shaded regions indicate 3σ (i.e. 3-standard deviation) intervals and ages are binned in 5-year intervals. The shown nationwide forecasts exhibit a remarkable similarity to the actual observations; despite a significant and unexpected rise in overdose fatalities in the pandemic year 2020 compared to 2019, the EnKF forecast of age-stratified overdose deaths for the year 2021 remains in close agreement with the reported number of fatalities. In the Materials and methods section, we compare our EnKF forecasts with those generated by two heuristics, which are based on the assumption that present trends persist in the future.

In addition to forecasting fatalities in each age group, we also used our EnKF to estimate the trajectories of $\hat{\mu}_d$, $\hat{a}_1^{max}$, $\hat{a}_2^{max}$, $\hat{r}_1$,

and $\hat{r}_2$ over the 1999–2021 interval. The EnKF was initialized with $\hat{\mu}_d = 0.2\%$ per year, consistent with 1999 data (66); we also set the initial values $\hat{a}_1^{\max} = 30$ years, $\hat{a}_2^{\max} = 45$ years, and $\hat{r}_1 = \hat{r}_2 = 2\%$ per year. Figure 1f shows that $\hat{\mu}_d$ increased 6-fold in the past 20 years, rising from about 0.2% per year to 1.2% per year. This finding is largely independent of the initial value of $\hat{\mu}_d$; the final estimate, $\hat{\mu}_d = 1.2\%$ overdose deaths per year is larger than the baseline Gompertz–Makeham–Siler approximation for all ages under 60 years old.

The trajectory of the quantities $\hat{a}_1^{\max}$, $\hat{a}_2^{\max}$ presented in Fig. 1g shows that between 1999 and 2009 these values remain within the 30–35 and 45–50 range, respectively. In later years however, while $\hat{a}_2^{\max}$ decreases only mildly, there is a strong descent of $\hat{a}_1^{\max}$ toward lower values, even below 20 years, indicating a substantial inflow of younger users. Figure 1g also shows that $\hat{r}_1$ increases within the period of observation and that in 2015 it surpasses $\hat{r}_2$, so that the influx of the population that uses high-risk drugs is dominated by the distribution peaked at the younger age $\hat{a}_1^{\max}$. The shift of the onset of drug use toward younger ages that is observed starting in 2015 is consistent with the concurrent emergence of high mortality rates within the 25–35 age group as seen in Fig. 1b.

Assessing the mortality rate of the population that uses high-risk drugs is challenging since a large number of subjects must be recruited and followed to evaluate frequency of rare events like death. Studies are typically conducted among formerly incarcerated persons, SUD patients enrolled in treatment clinical trials or who have been hospitalized and monitored postdischarge (59, 60). These studies reveal that people who use high-risk drugs exhibit elevated mortality compared to the general population primarily due to fatal overdoses, but also due to viral infections, cardiovascular disease, and cancer (67, 68). Overdose-specific mortality rates among persons who use high-risk drugs vary between 0.2 and 1 % depending on gender, drug of choice, treatment type (68–71); other metastudies reveal that the mortality rate for persons who use high-risk opioids is about 0.7 % (72). Our estimates for $\hat{\mu}_d$ from Fig. 1f are in agreement with these values; in addition, our work enables the tracking of longitudinal changes in overdose mortality rates throughout the entire 1999–2021 period, uncovering a significant and alarming surge in mortality rates among persons who use high-risk drugs. Particularly noteworthy is the pronounced increase observed during the pandemic years 2020 and 2021. To summarize, our findings reveal a staggering 6-fold increase in mortality rates among people at high risk of overdose, a generational shift toward drug use at younger ages, and alarming numbers of overdose deaths among individuals up to 30 years old. These results emphasize the necessity for more focused approaches in intervention and prevention strategies.

Forecasts for future years

In Fig. 2, we show nationwide forecasts for the years 2022, 2023, and 2024. The 2022 forecast is based on the 2021 values of $\hat{\mu}_d$, $\hat{a}_1^{\max}$, $\hat{a}_2^{\max}$, $\hat{r}_1$, $\hat{r}_2$. Since the final data for 2022 are not available at the time of writing, we incorporate provisional 2022 data in our data-assimilation cycle to determine forecasts for 2023 and 2024. The unfolding of the drug crisis over the past few years has been strongly influenced by the COVID-19 pandemic and the extraordinary rise of overdose deaths recorded in 2020 and 2021. This resulted in anomalies in the forecasts for 2020 and 2022, when the pandemic began receding. Our EnKF, however, was able to adjust quite efficiently to the sudden changes imparted by the pandemic. For example, while our 2020 forecast (not shown) was an undershoot compared to actual 2020 data, feeding the 2020 data into the EnKF led to good agreement between forecasts and observations for 2021 as can be seen in Fig. 1e. Similarly, implementing 2020 and 2021 data in the EnKF results in overshoots for the 2022 forecast compared to provisional 2022 data, as the effects of the pandemic began waning. However, once the 2022 provisional data were included, our EnKF forecasts for 2023 and 2024 adjusted to overdose counts that are closer to those recorded in 2021.

Our EnKF forecasts suggest that fatalities will remain largest among those younger than 30 and that drug-overdose counts will remain elevated for all relevant age groups.

County-level variation

Although the overall number of drug-overdose fatalities in the United States is rising, it varies significantly across jurisdictions. Previous analyses at the state level examined the impact of various factors on overdose deaths, including the availability of carfentanil (73), and demographic variables such as gender and race (13). County-level analyses have been conducted in various states, including California (74, 75), New York (76, 77), Ohio (78), and Michigan (79, 80). Furthermore, differences in overdose dynamics between urban and rural areas have been documented throughout the United States over the past two decades (81–83). Collectively, these studies highlight the need for region-specific interventions. Building upon the aforementioned works, we now overview quantitative variations in overdose mortality in the United States at the county level from 1999 to 2021.

Figure 3a shows the distribution of county-stratified drug-overdose fatalities for select years between 2000 and 2020. Only counties with statistically significant fatalities of at least 10 individuals per year are shown. The number of counties that reached this significance threshold increased from 61 counties (out of 3,147) in 1999 to 742 (out of 3,142) in 2021, as reported in the CDC WONDER database. Between 1999 and 2021, numerous counties reported annual numbers of overdose fatalities below 100. However, during the same period, the number of counties experiencing between 100 and 1,000 annual overdose fatalities steadily increased. In 2020 (blue disks), a few counties even recorded close to 1,000 overdose deaths. Crude rates, defined as the number of deaths per 100,000 persons, also increased significantly over time, as seen in Fig. 3b: in the year 2000, the mean crude rate among all counties for which data were available was 4.3 cases per 100,000, in 2020 it was 31.5 cases per 100,000. This 7-fold increase is consistent with the similar rise in $\hat{\mu}_d$ as inferred by our EnKF on the national level. The distributions in Fig. 3b also show that the crude rates exhibit a high degree of variability across counties.

Figure 3d shows that in 2000, only counties with populations larger than 100,000 residents experienced statistically significant numbers of drug-overdose fatalities. In the years since, crude rates substantially increased for these counties, especially between 2010 and 2020, while smaller counties with population sizes of about 10,000 also started reporting significant numbers of fatal overdoses, as shown in Fig. 3e. This indicates that the drug-overdose epidemic has permeated all jurisdictions, regardless of population.

The heat maps in Fig. 3c and f confirm the increase in the number of counties affected by the drug epidemic between 2000 and 2020. Notice that the scales of the two color bars differ by a factor of 10. Figure 3c shows that in 2000, the most affected areas were population centers in the Western United States: that year, the largest overdose fatality counts occurred in Maricopa County, AZ; Los Angeles County, CA; and Clark County, NV (Fig. 3g) and the largest crude rates were reported in Bernalillo County, NM; Washoe County, NV; and Salt Lake County, UT (Fig. 3h). In 2020,

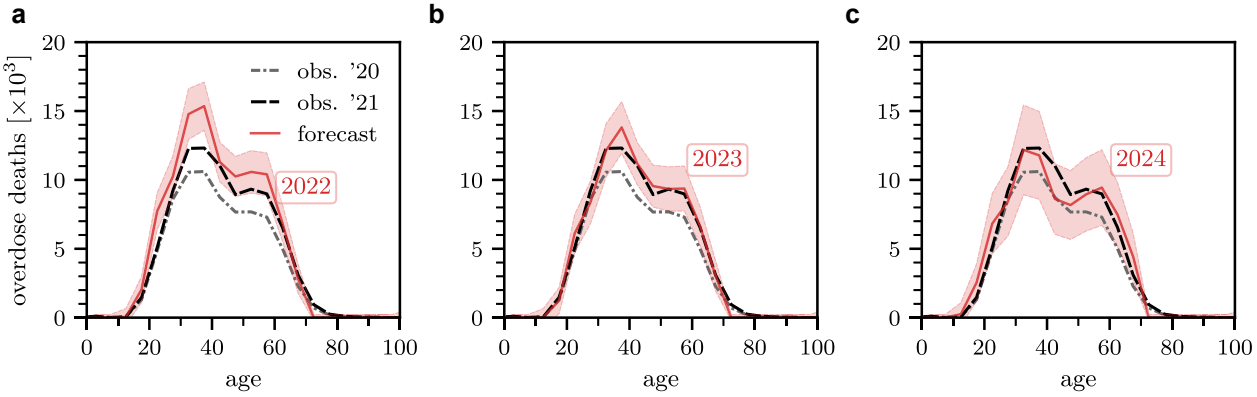

Fig. 2. Forecasting nationwide overdose fatalities for 2022–2024. a) Forecasts of overdose deaths in the United States as a function of age (0–100 years) for the year 2022. The forecast (solid red curve) is higher than 2020 and 2021 observations (dash-dotted curve gray and dashed black curve) due to the EnKF following the trend set by the pandemic years 2020 and 2021, both marked by large increases in overdose deaths. Provisional data for 2022 (not shown) suggests that overdose deaths in each age group are comparable to the 2021 counts. Light red-shaded regions show values within the 3σ range. b and c) Forecasts of overdose deaths in the United States as a function of age (0–100 years) for the years 2023 and 2024, using 2022 provisional data in the EnKF. The predicted number of overdose deaths will remain near the record highs of 2021. Due to the lack of observation data in 2023, the CIs increase in 2024 due to the larger uncertainty as time progresses. Ages are binned in 5-year intervals and prediction values are displayed at the center of each bin interval.

many regions in the Central and Eastern United States also became heavily impacted by the drug epidemic, including many smaller population counties (Fig. 3f). The largest 2020 fatality county were reported in Los Angeles County, CA; Cook County, IL; and Maricopa County, AZ (Fig. 3g). The largest 2020 crude rates were registered in Wyoming County, WV; McDowell County, WV; and Floyd County, KY (Fig. 3h).

In 2000, Los Angeles County, the most populous in the United States, accounted for approximately 12% of the total population of the 59 counties with statistically significant overdose fatalities. Its proportion of overdose fatalities was about 10%. However, due to more counties reporting large numbers of overdose deaths, by 2020, Los Angeles County's population represented only 4% of the total population of the 640 counties with statistically significant overdose deaths. That same year, Los Angeles County contributed to approximately 3% of the registered overdose deaths at the county level. Furthermore, in 2020, despite the 10 most populous counties in the United States being home to roughly 16% of the population (out of the 258 million associated with the 640 counties with statistically significant overdose fatalities), they recorded less than 13% of the number of overdose fatalities among these 640 counties. Conversely, the 50 least populous counties, with less than 1% of the total population among the 640 affected counties, accounted for more than 2% of the number of overdose deaths. These statistics highlight the shifting patterns of overdose fatalities, with a notable rise of overdose deaths in less populated counties. To better understand how the drug deaths are shared across counties, we consider the Gini coefficient (84, 85), a measure of inequality among a set of N_c values of a distribution; in our case, N_c is the number of counties that have reported statistically significant numbers of overdose deaths. We compute the Gini index by plotting the proportion of the total number of overdose fatalities accumulated across counties against the cumulative population fraction across counties. The lower bound for the Gini index is 0 (perfect equality, indicating that overdose deaths and county populations are proportional), and the upper bound is $1 - 1/N_c$ (perfect inequality, indicating that all overdose deaths occurred within a single county). We find that the Gini coefficient dropped from a value of about 0.2 in the year 2000 ($N_c = 59$), to about 0.07 in the year 2021 ($N_c = 640$), which is consistent with increases in the number of counties affected by the drug-overdose epidemic.

In the next section, we employ the modeling and forecasting techniques established in the previous section to examine the progression of age-specific overdose fatality counts in three specific regions: Los Angeles County, CA; Cook County, IL; and the combined area of New York City's five boroughs (The Bronx, Brooklyn, Manhattan, Queens, and Staten Island).

Forecasting overdose fatalities in three counties

The dynamics of overdose fatalities in Los Angeles County, Cook County, and the five boroughs of New York City have unfolded in substantially different ways over the past two decades. In 1999, Cook County reported a total of 14 overdose fatalities, New York City recorded 52, and Los Angeles County 384. By 2021, these numbers had risen to 1,688 overdose deaths for Cook County, 2,091 for Los Angeles County, and 2,124 for New York City. Cook County experienced the most striking rise in overdose fatalities between 1999 and 2021: an unprecedented 120-fold increase. In comparison, New York City experienced a 40-fold increase and Los Angeles County a 5-fold increase. These numbers are even more striking given Cook County's smaller population (5.1 million in 2021), compared to the population of Los Angeles County (9.7 million in 2021) and New York City (8.3 million in 2021). In the Materials and methods section, we provide further details on the trends in population, overdose mortality, and drug types across these three regions.

As in our nationwide forecasts, we use an EnKF in conjunction with an age-structured overdose mortality model that accounts for the underlying age variation in the county populations. Because of the relatively small overdose fatality counts in Cook County and New York City in the early 2000s, we do not report forecasts for 2001 as done nationwide and for Los Angeles County, but we use years from 2013 onwards for which enough data are available across all age groups. Specifically, in Cook County, the total number of reported age-stratified drug-overdose fatalities remained below 100 for most years prior to 2013. This points to a delayed emergence of the overdose death crisis in this county which became extraordinarily acute in just a few years. A finer analysis reveals that the largest increases in drug-overdose mortality in Cook County are due to heroin (a 10-fold rise between 2012 and 2013) and to fentanyl (a 5-fold rise between 2015 and 2016).

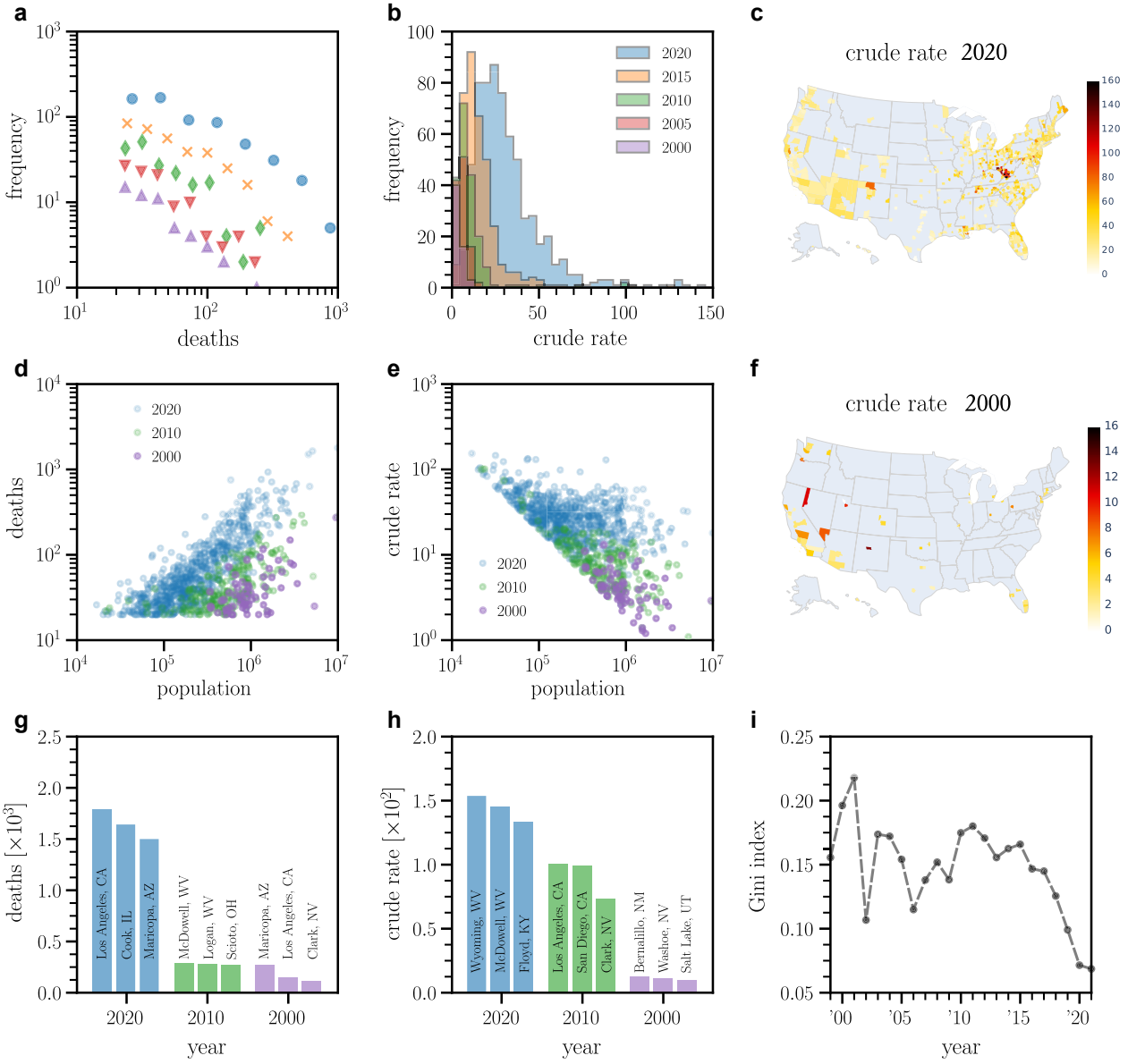

Fig. 3. Drug-overdose fatalities in the United States at the county level. a) Histogram of county-stratified drug-overdose fatalities for different years (blue disks: 2020, orange crosses: 2015, green diamonds: 2010, red inverted triangles: 2005, purple triangles: 2000). b) Histogram of county-stratified crude rate data for different years. d) County-stratified drug-overdose fatalities as a function of the corresponding county populations. e) County-stratified crude rates as a function of the corresponding county populations. c), f) Crude rates across different counties in 2020 and 2000. The scales of the two color bars differ by a factor of 10. In the gray regions, either no data or a statistically not significant number of cases were reported. In all panels, we did not include data for which at least one database entry (e.g. fatalities and crude rate) was marked unreliable. The minimum number of deaths in the remaining data is 20. Hence, the crude rate data in e) lies above of the minimum crude rate given by 2×10^6/Population. g) The three counties with the largest overdose death tolls in 2020, 2010, and 2000. h) The three counties with the largest overdose crude rates in 2020, 2010, and 2000. i) The Gini index across different years. A Gini index of 0 means that the crude rate is the same across all counties. If all overdose fatalities were concentrated in one out of N_c counties, the Gini index would be $1 - 1/N_c$. The number of counties with statistically significant fatality counts (>10 deaths in a given year) and crude rates are $N_c = 61$ in 1999 and $N_c = 742$ in 2021.

We show our results in Fig. 4. The first three historical forecasts for each region are compared to the corresponding observational data (dashed black curves). Since the observational data for 2023 are not available at the time of writing, we display the 2020 data (dash-dotted gray curve) and the 2021 data (dashed black curve) in the 2023 forecast. As done for the nationwide forecasts, we incorporate provisional 2022 data in our data-assimilation cycle to determine the 2023 forecasts at the county level. For Los Angeles County, we initialized the EnKF with $\hat{\mu}_d = 0.25\%$ per year and $\hat{r}_1 = \hat{r}_2 = 2\%$ per year. The simulations for Cook County and New York City start in 2013, and we initially set $\hat{\mu}_d = 0.5\%$ per year and $\hat{r}_1 = \hat{r}_2 = 6\%$ per year. The estimated drug-induced mortality rate for Cook County and New York City in 2021 are, respectively, $\hat{\mu}_d = 1.9\%$ per year and $\hat{\mu}_d = 1.1\%$ per year, substantially larger than the 2021 estimates for Los Angeles County ($\hat{\mu}_d = 0.5\%$ per year). This is also consistent with the much larger increase in overdose fatalities in Cook County and New York City compared to Los Angeles County.

Discussion

Escalating drug-induced deaths have been a major public-health challenge in the United States for more than a century. The

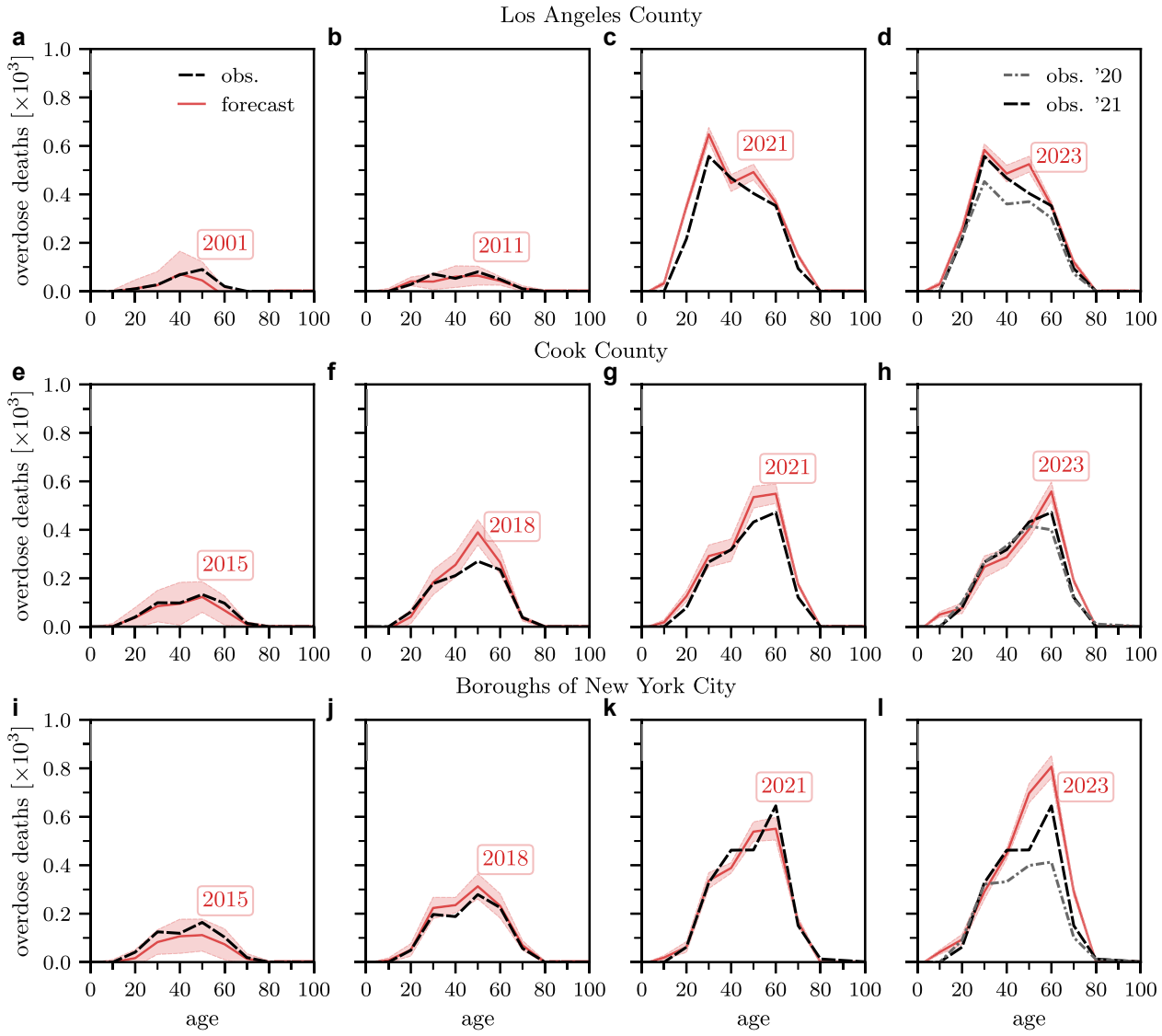

Fig. 4. Forecasting overdose fatalities in select United States jurisdictions. a–d) Forecasts of overdose deaths in Los Angeles County as a function of age (0100 years) for 2001, 2011, 2021, and 2023. Solid red curves and shaded regions indicate mean predictions and 3σ intervals, respectively. Observed fatalities are indicated by dashed black curves. Because observational data for 2023 are not available at the time of writing, we show 2020 data (dash-dotted gray curve) and 2021 data (dashed black curve) in panels d, h, and l). e–h) Forecasts for Cook County using the same graphic representations as in panels a–d and for the years 2015, 2018, 2021, 2023. i–l) Forecasts for the five boroughs that comprise New York City (The Bronx, Brooklyn, Manhattan, Queens, and Staten Island) using the same graphic representations as in panels a–d and for the years 2007, 2014, 2021, 2023. Although the numerical escalation in drug-overdose deaths in Cook County (2015–2021) and New York City (2007–2021) is similar to what is observed for Los Angeles County, the timelines are much accelerated. In 2021, the population of Los Angeles County was 9.7 million, in Cook County 5.1 million, and in New York City 8.3 million, indicating a more acute crisis in Cook County. Ages are binned in 10-year intervals and prediction values are displayed at the center of each bin interval.

overprescription of morphine and opium led to an epidemic that affected almost 5 in 1,000 Americans in the 1890s (86). This widespread crisis spurred a number of acts and regulations in the early 20th century that succeeded in lowering opiate use and mortality rates (87, 88). The current epidemic involves a significantly higher prevalence of high-risk drug use and has unfolded via distinct spatiotemporal mortality waves driven by different drug types and localized subepidemics. Being able to forecast the complex evolution of fatal drug overdoses at the national, regional, or county levels, would represent major advancements in helping curb high-risk drug use. In this work, we developed a forecasting method that combines an age-structured model of addiction and overdose mortality with observational data derived from the CDC WONDER database through a data-assimilation approach. By applying our method to nationwide data as well as to three representative areas (Los Angeles County, Cook County, and the five boroughs of New York City), we showed its ability to provide near-term forecasts, to extract epidemiological parameters, and to capture the heterogeneity in overdose mortality across different counties. Since the demographics and geography of high-risk drug use are in constant flux, we believe our data-assimilation approach holds promise for informing targeted prevention and preparedness interventions aimed at curbing drug-overdose deaths.

The nationwide drug-induced mortality rate has risen almost 6-fold in the past two decades, surpassing 1.1% per year among persons using high-risk drugs. This rate exceeds the baseline Gompertz–Makeham–Siler mortality rate for all groups under 60 years old. Our county-level analysis reveals significant variations in overdose fatality trends. For example, although at the onset of

1999–2021 period, Los Angeles County had a higher drug-induced mortality rate than Cook County and New York City, in 2021 it had the lowest, at 0.5% per year, compared to Cook County at 1.9% per year and New York City at 1.1 % per year. Furthermore, the annual number of overdose deaths in Cook County and New York City grew much faster than in Los Angeles County in recent years. This points to a delayed, yet severe growth of drug overdoses in New York City and especially in Cook County. For past years, our predictions are in good agreement with tallied data. For the year 2023, we predict relatively stable levels of drug overdoses compared to the pandemic year 2021, both nationwide and in Los Angeles County and Cook County, based on provisional 2022 data. For New York City, our forecasts indicate a further increase of overdose deaths. Specifically, we expect drug-overdose deaths to slightly increase nationwide by 1 % compared to the values recorded in 2021, by 12% in Los Angeles County, by 26% in New York City but only by 7% in Cook County. Prior work (14, 73) has shown that the number of overdose deaths in the United States has closely followed an exponential growth pattern over the past four decades. The substantial increase in overdose deaths between 2019 and 2021 is likely linked to the COVID-19 pandemic. Provisional data indicate that, as the pandemic began receding, the total number of overdose deaths in 2022 is mostly comparable to the 2021 counts, albeit slightly larger.

We also find that the overdose epidemic has spread to more counties over time. In the year 1999, 61 counties out of 3,147 had statistically significant overdose fatalities, whereas this number rose to 742 out of 3,142 in 2021. Not only has the number of affected counties grown, but their relative contributions to the overall overdose fatality count have become more evenly distributed over the years. This finding implies that managing the overdose epidemic cannot be simply accomplished by targeting a few specific counties, rather each jurisdiction must develop specific plans tailored to their unique sociodemographic and economic profiles.

Several limitations of this study are noteworthy. Our findings are based on four drug categories with the highest crude rates available in the CDC WONDER database: fentanyl (T40.4), prescription opioids (T40.2), heroin (T40.1), and methamphetamines (T43.6). We did not include other categories such as T40.3 (methadone) or T40.5 (cocaine) in our analysis due to their lower mortality rates. The dynamics of fatalities associated with these drug categories may differ from the fatality trends observed in our analysis. Furthermore, in certain jurisdictions, fatality data are unavailable as the CDC WONDER portal suppresses entries where the number of deaths is less than 10. Additionally, some overdose cases may involve multiple drugs. In such instances, deaths are counted in all relevant categories, resulting in multiple counts. Finally, comparisons of opioid-related overdose death rates at the national, state, and county levels may be influenced by significant variations in the reporting of specific drugs involved in overdose deaths. Changes in drug reporting specificity over time and across different states and counties can lead to potentially misleading conclusions regarding actual drug-specific death rates (89).

There are several potential avenues for future work. Although we only analyzed three large urban areas, our method can be applied to other jurisdictions and/or to forecast regional drug-overdose mortality by gender, race or drug type; the resulting projections may help to guide more targeted intervention efforts. In less populated areas, the number of fatalities may not be sufficiently large for a meaningful age-stratified analysis; similarly for specific gender, race or drug-type categories. In these cases, pooling data from several neighboring jurisdictions with similar socioeconomic characteristics, using larger age-binning or considering biannual forecasts may yield more meaningful results. Alternatively, for small-number cases, a stochastic version of the Kermack–McKendrick model may be used as to evolve the state variables probabilistically (90). Large deviations of observed data from our projections would signal fundamental changes to the illicit drug landscape in the form of effective prevention and treatment programs, or in the consumption of more addictive or lethal substances. On the mathematical modeling side, our Kermack–McKendrick model may be expanded to include shifts among drug types, or to distinguish whether overdoses occur among those with SUD or first time, occasional users. On the data side, another potential opportunity for future work is integrating our forecasting method into an ensemble model to leverage different strengths for improved forecasting accuracy. Finally, one may also study the numerical stability and forecasting accuracy of alternative ensemble-based Kalman filters, such as ensemble adjustment Kalman filters (91), or incorporate backward passes and smoothing techniques into our method to potentially enhance earlier parameter estimates (92).

Materials and methods

Age-structured overdose model

The mathematical model we use to describe the age-stratified evolution of the population that uses high-risk drugs is given by

$$\left[\frac{\partial}{\partial a}+\frac{\partial}{\partial t}\right]n(a,\,t)=-\mu(a,\,t)n(a,\,t)+r(a,\,t)[N(a,\,t)-n(a,\,t)], \qquad (1)$$

where $n(a,\,t)da$ is the drug-using population with age between a and $a+da$ at time t. The associated mortality is $\mu(a,\,t)$ and $r(a,\,t)$ is the influx rate of new users from $N(a,\,t)-n(a,\,t)$, the pool of individuals not engaged in high-risk substance use. Finally, $N(a,\,t)$ is the general population with age between a and $a+da$ at time t. We set the initial age and time $a_0=t_0=0$. The initial distribution of SUD cases is given by $n(a,\,t=0)=\rho(a)$. We also set $n(a=0,\,t)=0$ such that no population of age $a=0$ exists at any time. We solve Eq. 1 using the method of characteristics and distinguish the two cases $a\geq t$ and $a<t$. For $a\geq t$, the characteristic begins at $t=0$ and $n(a,\,t)$ will remain constant along $a=t$, yielding

$$n(a,\,t)=\rho(a-t)e^{-\int_0^t\mu(s+a-t,s)+r(s+a-t,s)\,ds} + \int_0^t r(s+a-t,\,s)N(s+a-t,\,s)e^{-\int_s^t\mu(z+a-t,z)+r(z+a-t,z)dz}\,ds \quad (a\geq t). \qquad (2)$$

For $a<t$, the characteristic will begin at $a=0$ and $n(a,\,t)$ will remain constant along $t=a$ so that

$$n(a,\,t)=\int_0^a r(s,\,s-a+t)N(s,\,s-a+t)e^{-\int_s^a\mu(z,z-a+t)+r(z,z-a+t)\,dz}\,ds \quad (a<t). \qquad (3)$$

We write the mortality rate $\mu(a,\,t)$ as the sum of a baseline mortality rate, $\mu_0(a,\,t)$, and a drug-caused excess mortality rate, $\mu_d(a,\,t)$, so that $\mu(a,\,t)=\mu_0(a,\,t)+\mu_d(a,\,t)$. In principle, $\mu_0(a,\,t)$ could be derived from records of yearly mortality data. However, this approach would make our numerical computations very time-consuming, so instead we use the Gompertz–Makeham–Siler mortality model for human death (61–65) as an approximation and assume $\mu_0(a,\,t)=\mu_0(a)$ to be time-independent. Finally, the quantity $\mu_d(a,\,t)=\mu_d$ is assumed to be age-independent and inferred from data, so that we effectively neglect any time dependence over a data-assimilation cycle of 1 year. Within a data-assimilation window of 1 year, we thus set

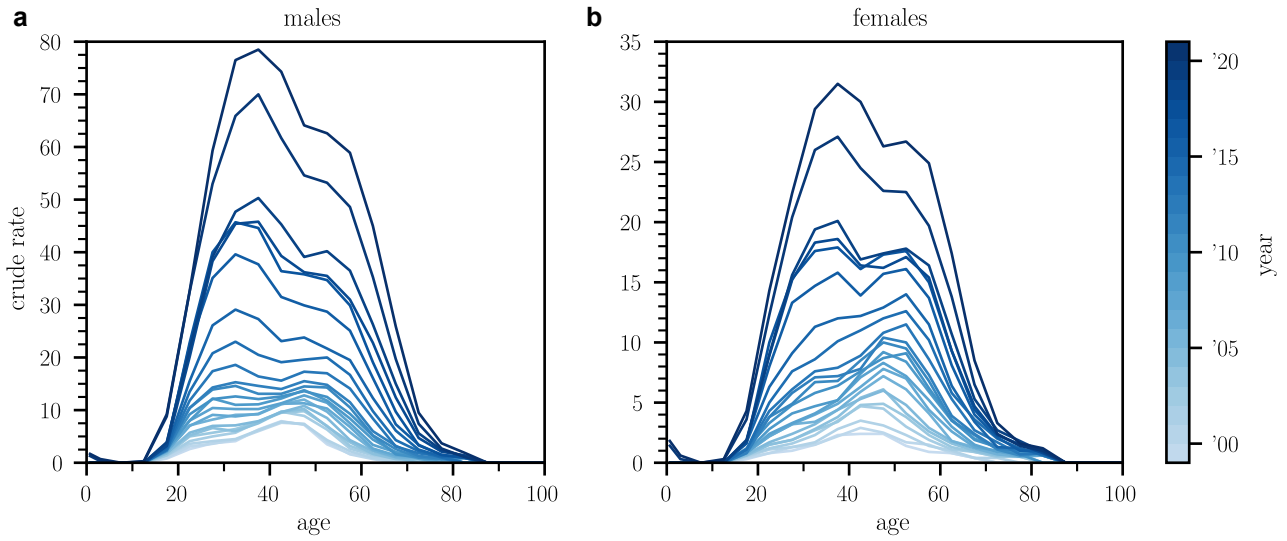

Fig. 5. Age-structured overdose crude rates in the United States. Age-structured overdose crude rates in the United States, distinguished by gender: a) males and b) females.

$$\mu(a,t)=\mu_0(a)+\mu_{\rm d}=\gamma_1 e^{-\lambda_1 a}+\gamma_2+\lambda_2 e^{\lambda_2(a-M)}+\mu_{\rm d}\quad(\gamma_{1,2},\lambda_{1,2},M>0), \tag{4}$$

where $\gamma_1=0.00258$/year, $\gamma_2=0.00037$/year, $\lambda_1=5.09657$/year, $\lambda_2=0.09040$/year, and $M=83.22956$ year in accordance with Ref. (65). Since the majority of drug-overdose fatalities occur among males (13, 93) as seen in Fig. 5, we explicitly used parameters pertaining to males in the United States. The parameters in Eq. 4 vary slightly from year to year, so we selected the ones for 2010. It is worth noting that we conducted a sensitivity analysis through additional simulations using variations on the above choices for γ_1, γ_2, λ_1, λ_2, M including values specific to females in the United States. We find that the choice of these parameters does not substantially affect our results.

We allow the parameter $\mu_{\rm d}$ to change from one year to the next. This choice is dictated by the CDC WONDER database providing yearly lists of overdose deaths, although monthly updates could also be implemented. To describe an age-dependent influx into the pool of persons who use high-risk drugs, we set $r(a,t)\equiv r(a)=\left[r_1 f(a;\alpha_1,\beta_1)+r_2 f(a;\alpha_2,\beta_2)\right]/2$, where $f(a;\alpha,\beta)=\beta^\alpha/\Gamma(\alpha)\,a^{\alpha-1}e^{-\beta a}$ is the gamma distribution with shape and rate parameters α and β. The maximum of a gamma function is given at the age $a^{\max}=(\alpha-1)/\beta$. The evolution of $n(a,t)$ in data-assimilation cycles requires us to evaluate the derivative of Eq. 2 with respect to (w.r.t.) t. We will use a superscript $'$ to denote differentiation of a function w.r.t. its first or only argument and a subscript $'$ to denote differentiation w.r.t. to a functions second argument. For $a\geq t$, the rate of change of $n(a,t)$ can thus be written as

$$\begin{aligned}\frac{\partial n(a,t)}{\partial t}=&-\Big(\rho'(a-t)+\rho(a-t)\left[\gamma_1 e^{-\lambda_1(a-t)}+\gamma_2+\lambda_2 e^{\lambda_2(a-t-M)}+\mu_{\rm d}+r(a-t)\right]\Big)\\&\times e^{\gamma_1 e^{-\lambda_1 a}(1-e^{\lambda_1 t})/\lambda_1-e^{\lambda_2(a-M)}(1-e^{-\lambda_2 t})-(\gamma_2+\mu_{\rm d})t}\,e^{-\int_0^t r(s+a-t)\,{\rm d}s}+r(a)N(a,t)\\&-\int_0^t e^{\gamma_1 e^{-\lambda_1 a}(1-e^{\lambda_1(t-s)})/\lambda_1-e^{\lambda_2(a-M)}(1-e^{-\lambda_2(t-s)})-(\gamma_2+\mu_{\rm d})(t-s)}\,e^{-\int_s^t r(z+a-t)\,{\rm d}z}\\&\times N(s+a-t,s)\,r(s+a-t)\Big(\left[\gamma_1 e^{-\lambda_1(s+a-t)}+\gamma_2+\lambda_2 e^{\lambda_2(s+a-t-M)}\right.\\&\left.+\mu_{\rm d}+r(s+a-t)\right]+r'(s+a-t)\Big)\,{\rm d}s\\&-\int_0^t e^{\gamma_1 e^{-\lambda_1 a}(1-e^{\lambda_1(t-s)})/\lambda_1-e^{\lambda_2(a-M)}(1-e^{-\lambda_2(t-s)})-(\gamma_2+\mu_{\rm d})(t-s)}\,e^{-\int_s^t r(z+a-t)\,{\rm d}z}\\&\times N'(s+a-t,s)r(s+a-t)\,{\rm d}s.\end{aligned} \tag{5}$$

For $a<t$, we obtain

$$\begin{aligned}\frac{\partial n(a,t)}{\partial t}=&\int_0^a r(s)N_r(s,s-a+t)\\&\times e^{\gamma_1(e^{-\lambda_1 a}-e^{-\lambda_1 s})/\lambda_1-e^{-\lambda_2 M}(e^{\lambda_2 a}-e^{\lambda_2 s})-(\gamma_2+\mu_{\rm d})(a-s)}\,e^{-\int_s^a r(z)\,{\rm d}z}\,{\rm d}s.\end{aligned} \tag{6}$$

The integrals $\int_s^t r(z+a-t)\,{\rm d}z$, $\int_0^t r(s+a-t)\,{\rm d}s$, and $\int_s^a r(z)\,{\rm d}z$ can be evaluated using the identity

$$\int_s^t \frac{\beta^\alpha}{\Gamma(\alpha)}(z+a-t)^{\alpha-1}e^{-\beta(z+a-t)}\,{\rm d}z=\frac{1}{\Gamma(\alpha)}\Big[\Gamma(\alpha,(a-t+s)\beta)-\Gamma(\alpha,a\beta)\Big], \tag{7}$$

where $\Gamma(s,x)=\int_x^\infty t^{s-1}e^{-t}\,{\rm d}t$ denotes the upper incomplete gamma function. We evaluate the remaining integrals $\int_0^t(\cdot)\,{\rm d}s$ numerically. Finally, the initial condition used to solve Eq. 1 and to obtain the simulation results in Figs. 1, 2, and 4 is

$$\rho(a)=0.015N_0 f(a;\alpha_0,\beta_0). \tag{8}$$

To obtain the curves shown in Figs. 1 and 2, we set $N_0=274{,}886{,}150$, the population of the United States between ages 0 and 85 in 1999. We also select $f(a;\alpha_0,\beta_0)$ to be a gamma distribution with shape and rate parameters α_0 and β_0, chosen as $\alpha_0=12$ and $\beta_0=1/(3\ {\rm year})$ such that the maximum of the distribution is at $a_{\max}=33$ years. The prefactor of 0.015 is chosen such that initially 1.5% of the population are using high-risk drugs consistent with corresponding survey data (66, 94). To obtain the curves in the county-level analysis, we initially set N_0 to match the respective population sizes between ages 0 and 85 for the years 1999 (Los Angeles County) and 2013 (Cook County and New York City). Specifically, N_0 was set to 9,437,290 for Los Angeles County, 8,405,837 for New York City, and 5,240,700 for Cook County. We also set $\alpha=17$ (Los Angeles County) and $\alpha=12$ (Cook County and New York City) and used the same value of β as in the national analysis.

Interpolating population data

We infer the age-structured population function $N(a,t)$ from nationwide population data that are available from the CDC WONDER database. In Fig. 6a, we show an interpolated and differentiable population function $N(a,t)$. In all simulations, we use interpolations that are based on bivariate splines of degree 2. In

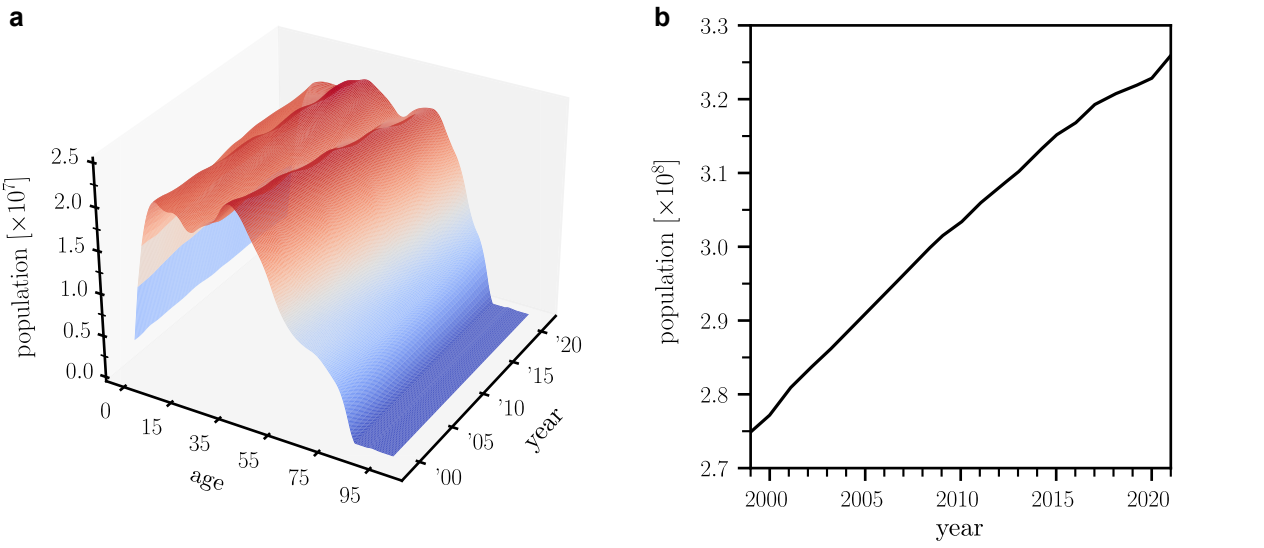

Fig. 6. Age-structured United States population data. a) Interpolated age-structured population data. The interpolation is based on bivariate splines of degree 2. b) The overall United States population increases almost linearly between 1999 and 2021.

Fig. 6b, we show the almost linear increase of the population of the United States between ages 0 and 85 from 1999 to 2021.

Ensemble Kalman filter

To combine the age-structured drug-overdose model (1) with corresponding observational data, we use an EnKF (58) as implemented in Refs. (33, 96). Figure 7 shows a schematic of the main EnKF steps.

In accordance with Refs. (31, 97), the evolution of any system state $\mathbf{x}(t)$ (e.g. number of SUD cases, mortality and addiction rates) and observed state $\mathbf{z}(t)$ (e.g. number of overdose fatalities) are described by the stochastic differential equations

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, t) + \mathbf{w}(t) \quad \mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(t)) \\ \mathbf{z} &= \mathbf{h}(\mathbf{x}, t) + \mathbf{v}(t) \quad \mathbf{v}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}(t)), \end{aligned} \tag{9}$$

where $\mathbf{Q}(t)$ and $\mathbf{R}(t)$ denote the covariance matrices associated with the Gaussian process noise $\mathcal{N}(\mathbf{0}, \mathbf{Q}(t))$ and Gaussian measurement noise $\mathcal{N}(\mathbf{0}, \mathbf{R}(t))$ at time t, respectively. We assume the quantities $\mathbf{Q}(t)$ and $\mathbf{R}(t)$ to be given. The function $\mathbf{f}(\cdot)$ describes the dynamics of the system state $\mathbf{x}(t)$, while $\mathbf{h}(\cdot)$ maps $\mathbf{x}(t)$ to a measurable quantity. Both functions can be nonlinear.

For the specific case of our age-structured model defined in Eq. 1, element $x_j(t)$ of the state vector $\mathbf{x}(t)$ corresponds to $n(a_j, t) \equiv n(a_0 + (j-1)\Delta a, t)$ ($j \in \{1, \ldots, N_a\}$), the density of individuals whose age lies within the $[a_0 + (j-1)\Delta a, a_0 + j\Delta a)$ interval at time t. Here, N_a and Δa denote the number of discretizations of the age interval and the corresponding age discretization step, respectively. Thus, we write

$$\mathbf{x}(t) = [n(a_1, t), n(a_2, t), \ldots]^\mathsf{T}. \tag{10}$$

For the numerical solution of Eq. 9, we also discretize the simulation time interval $[0, T]$ into N_t equidistant intervals of duration $\Delta t = T/N_t$. In all of our simulations, we fixed $\Delta t = 0.1$. To combine the mechanistic model in Eq. 1 with empirical data on overdose fatalities, we augment the system state (10) by

$$\tilde{D}(a_j, t) = \int_0^t \mu_d(\tilde{t}) n(a_j, \tilde{t})\, d\tilde{t} \quad (j \in \{1, \ldots, N_a\}), \tag{11}$$

where $\tilde{D}(a_j, t)$ is the cumulative number of overdose deaths in the age interval $[a_{j-1}, a_j)$ up to time t. To avoid dealing with large differences between predicted and observed fatalities in our numerical calculations, we normalize both quantities by dividing them by 1,000, thereby measuring overdose deaths per 1,000 individuals in the system state. Since we wish to estimate model parameters such as μ_d and $r_1, r_2, \alpha_1, \beta_1, \alpha_2, \beta_2$, we also augment the system state (10) by the log-transforms $\tilde{\mu}_d = \log(\mu_d)$, $\tilde{r}_1 = \log(r_1)$, $\tilde{r}_2 = \log(r_2)$, $\tilde{\alpha}_1 = \log(\alpha_1)$, $\tilde{\beta}_1 = \log(\beta)_1$, $\tilde{\alpha}_2 = \log(\alpha_2)$, and $\tilde{\beta}_2 = \log(\beta_2)$. Therefore, the final augmented system state is

$$\mathbf{x}(t) = \left[n(a_1, t), \ldots, n(a_{N_a}, t), \tilde{D}(a_1, t), \ldots, \tilde{D}(a_{N_a}, t), \tilde{\mu}_d, \tilde{r}_1, \tilde{r}_2, \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_2, \tilde{\beta}_2\right]^\mathsf{T}. \tag{12}$$

Prior to each prediction step, we apply an exponential transform to render the parameters μ_d and $r_1, r_2, \alpha_1, \beta_1, \alpha_2, \beta_2$ positive and avoid sign changes. To accurately solve the evolution of $n(a, t)$ numerically, we must use a sufficiently large number of age windows N_a in our simulations. However, since the age windows in our simulations are more granular than those available from overdose fatality data, we apply a coarse-graining procedure. The nationwide CDC WONDER data we utilized have 22 age groups with $a'_0 = 0$, $a'_{22} = 120$ and $\Delta a'_1 = 1, \Delta a'_2 = 4, \Delta a'_3 = 5, \ldots, \Delta a'_{21} = 5, \Delta a'_{22} = 20$ years. To differentiate between the age discretization in the observational data and the age discretization in the underlying model, we employ a superscript $'$ notation. In the county-level forecasts, we employed 10-year age groups, effectively reducing the number of age groups from 22 to 11. The following discussion of the EnKF parameterization and implementation will be based on 22 age groups, but the same considerations also apply to the county-level analysis with fewer age groups.

To reduce granularity and combine the modeled quantities $\tilde{D}(a_j, t)$ with corresponding observational data, we numerically integrate $\tilde{D}(a_j, t)$ over the age windows $[a'_{\ell-1}, a'_\ell)$ ($\ell \in \{1, \ldots, 22\}$) to obtain the cumulative number of deaths $D(a'_\ell, t)$ in this age interval at time t. Here, $a'_\ell = a'_0 + \sum_{m=1}^{\ell} \Delta a_m$ for $\ell \geq 1$. Based on the described mapping of $\tilde{D}(a_j, t)$ to $D(a'_\ell, t)$, the measurement function becomes

$$\mathbf{h}(\mathbf{x}(t)) = [D(a'_1, t), D(a'_2, t), \ldots]^\mathsf{T}. \tag{13}$$

In our simulations, we set the initial values $n(a_j, 0) = \tilde{D}(a_j, 0) = 0$. In the nationwide analysis, we initially set $\mu_d = 2 \times 10^{-3}$/year, $r_1 = r_2 = 2 \times 10^{-2}$/year, $\alpha_1 = 10$, $\beta_1 = 1/(3\text{ year})$, $\alpha_2 = 15$, and $\beta_2 = 1/(3\text{ year})$. We have chosen the initial values of r_1, r_2 in accordance with corresponding empirical data on the number of

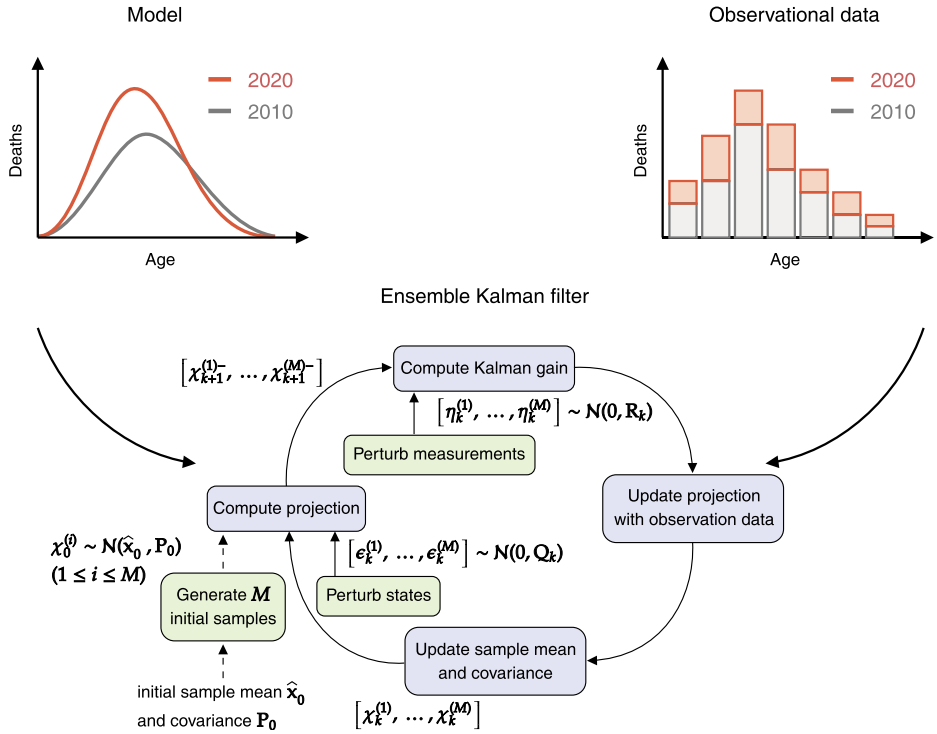

Fig. 7. EnKF schematic. We use an EnKF to combine a mechanistic model of drug-overdose fatalities with corresponding observational data. Blue boxes show the main steps (i.e. projection and update) in an EnKF cycle. Green boxes represent the initial sample generation process and perturbations that are added during the projection and update steps. The schematic is adapted from Ref. (95).

substance initiates (66, 94) The initial maximum values of the gamma distributions $f(a;\alpha_1,\beta_1), f(a;\alpha_2,\beta_2)$ are attained at ages $a_1^{\max} = (\alpha_1 - 1)/\beta_1 = 27$ years and $a_2^{\max} = (\alpha_2 - 1)/\beta_2 = 42$ years, respectively. All initial covariances are set to 10^{-4}, except for the diagonal elements associated with the log-transforms of μ_d, r_1, r_2 and $\alpha_1, \beta_1, \alpha_2, \beta_2$, which are set to 1, respectively. Process and observation noise covariances are assumed to be time-independent and given by $\mathbf{Q} = 10^{-4} J_{2N_a+7}$ and $\mathbf{R} = \mathrm{diag}(2\times10^{-3}, \ldots, 2\times10^{-3})$, respectively. Here, J_n denotes the $n \times n$ matrix of ones. Our simulations run from the beginning of 1999 until the end of 2024. The age discretization is $\Delta a = 1.2$ years with $a_0 = 0$ and $a_{N_a} = 120$ years. Population data are available for ages between 0 and 85. However, since overdose fatalities in groups below 10 years and above 70 years are statistically insignificant, we truncated the system state accordingly. To align model parameters with initial observational data, we performed two full data-assimilation cycles for the first year (1999) before starting the main forecasting algorithm that produces forecasts for all years from 1999 to 2024. The number of EnKF ensemble members is $M = 10^3$.

We used different initial values for the county-level forecasts. For Los Angeles County, the initial values were set as follows: $\mu_d = 2.5\times10^{-3}$/year, $r_1 = r_2 = 2\times10^{-2}$/year, and $\alpha_1 = \alpha_2 = 17$. For Cook County, the values were set as $\mu_d = 5\times10^{-3}$/year, $r_1 = r_2 = 6\times10^{-2}$/year, $\alpha_1 = 8$, and $\alpha_2 = 15$. Lastly, for New York City, the values were set as $\mu_d = 5\times10^{-3}$/year, $r_1 = r_2 = 6\times10^{-2}$/year, $\alpha_1 = 8$, and $\alpha_2 = 17$. The initial values of β_1 and β_2 are as in the nationwide analysis. To account for the smaller number of overdose fatalities at the county level, we adjust the process and observational noise matrix elements to have values of the order 10^{-8}– 10^{-7}. The number of EnKF ensemble members is $M = 10^2$ for all counties.

At every time point t, we use the EnKF to determine the state posterior distribution given all prior observations. Before starting the data-assimilation procedure, we generate an initial ensemble $[\chi_0^{(1)}, \ldots, \chi_0^{(M)}]$ that consists of M ensemble members $\chi_0^{(i)} \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ ($i \in \{1, \ldots, M\}$). The quantities $\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$ denote the given initial state and covariance estimates, respectively.

To perform forecast and update iterations using a Kalman filter, one uses state estimates $\chi_k^{(i)}$ at time t_k to calculate predicted state estimates $\chi_{k+1}^{(i)-}$ at time t_{k+1}. These predicted state estimates are then combined with observational data to obtain an updated state $\chi_{k+1}^{(i)}$. We use the superscript "–" in $\chi_{k+1}^{(i)-}$ to distinguish between predicted (i.e. prior) state estimates and updated (i.e. posterior) state estimates. In the remainder of this section, we describe the two main EnKF steps: (i) forecasting the evolution of the system state and (ii) updating the predicted state estimates using observational data. We use the shorthand notation $y_k \equiv y(t_k)$ to refer to a quantity y at time $t_k = k\Delta t$ ($k \in \{0, \ldots, N_t\}$).

(i) *Forecast step:* For each ensemble member, the predicted state estimate $\chi_{k+1}^{(i)-}$ at time t_{k+1} is given by

$$\chi_{k+1}^{(i)-} = \chi_k^{(i)} + \Delta t\, \mathbf{f}(\chi_k^{(i)}, t_k) + \epsilon_k^{(i)}, \tag{14}$$

where $\epsilon_k^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ models Gaussian process noise. Using the predicted state estimates $\hat{\mathbf{x}}_{k+1}^-$, we compute the corresponding ensemble mean, $\hat{\mathbf{x}}_{k+1}$, and covariance matrix, $(\mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^-)_{k+1}$, according to

$$\hat{\mathbf{x}}_{k+1}^- = \frac{1}{M}\sum_{i=1}^{M} \chi_{k+1}^{(i)-}, \tag{15}$$

$$(\mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^-)_{k+1} = \frac{1}{M-1}\sum_{i=1}^{M}\left[\chi_{k+1}^{(i)-} - \hat{\mathbf{x}}_{k+1}^-\right]\left[\chi_{k+1}^{(i)-} - \hat{\mathbf{x}}_{k+1}^-\right]^{\mathsf{T}}. \tag{16}$$

Although the covariance matrix $(\mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^-)_{k+1}$ is not required in the EnKF iteration, it is useful to estimate CIs of $\hat{\mathbf{x}}_{k+1}^-$.

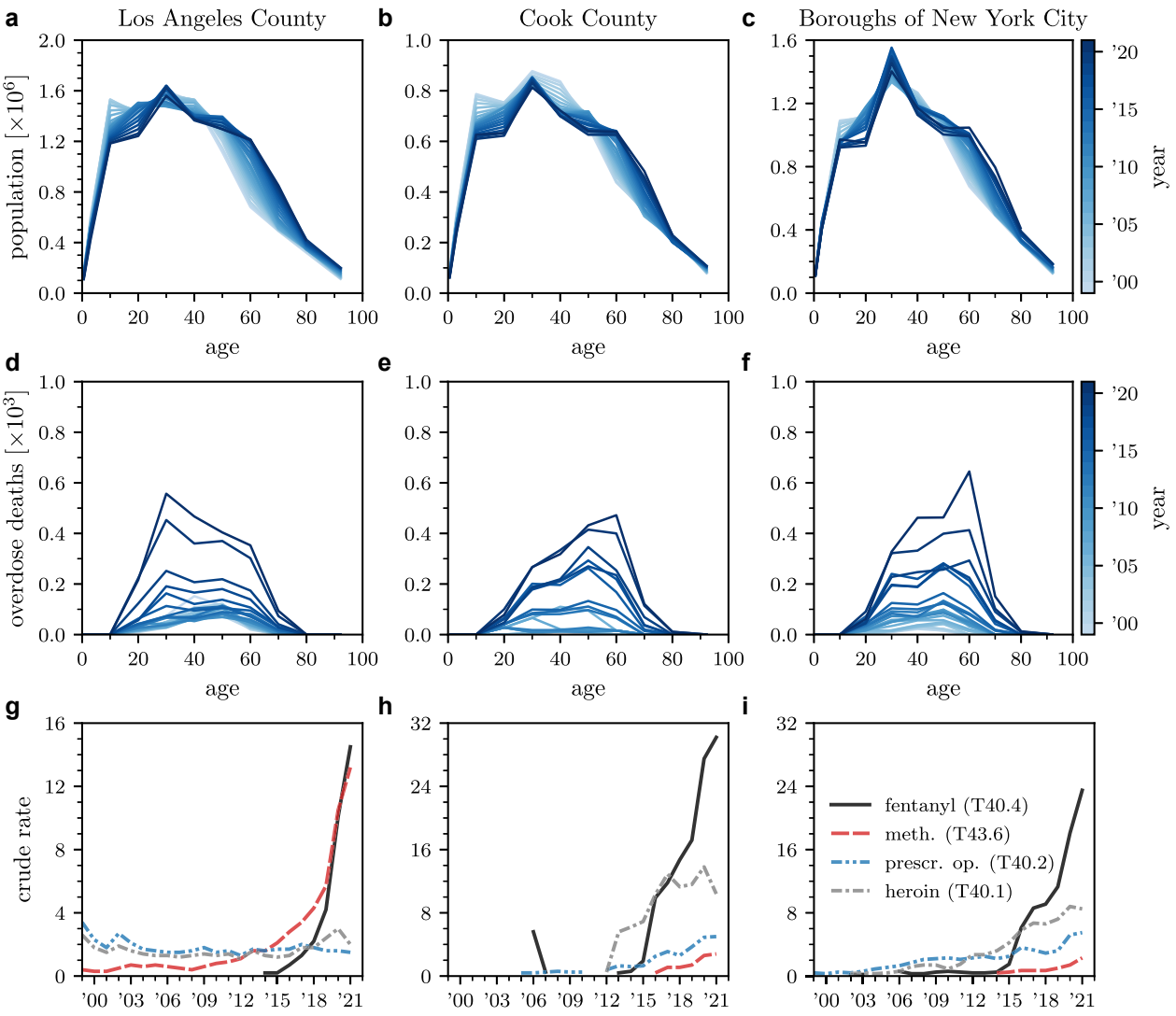

Fig. 8. Population, overdose, and drug-type trends in Los Angeles County, Cook County, and New York City. a–c) The population distribution across age groups (0–100 years) and time (1999–2021) in Los Angeles County, Cook County, and the five boroughs of New York City, with ages grouped in 10-year intervals. d–f) Overdose deaths in the same regions as a function of age (0–100 years) and time (1999–2021). g–i) Crude rates associated with fentanyl (T40.4), methamphetamines (T43.6), prescription opioids (T40.2), and heroin (T40.1) in Los Angeles County, Cook County, and the five boroughs of New York City from 1999 to 2021.

(ii) *Update step:* We first compute the ensemble mean of the predicted observation

$$\hat{\mathbf{z}}^{-}_{k+1} \equiv \frac{1}{M}\sum_{i=1}^{M}\mathbf{z}^{(i)-}_{k+1} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{h}(\boldsymbol{\chi}^{(i)-}_{k+1}) \tag{17}$$

as well as the corresponding covariances

$$\begin{aligned}(\mathbf{P}^{-}_{\hat{\mathbf{z}}\hat{\mathbf{z}}})_{k+1} &= \frac{1}{M-1}\sum_{i=1}^{M}\left[\mathbf{h}(\boldsymbol{\chi}^{(i)-}_{k+1}) - \hat{\mathbf{z}}^{-}_{k+1}\right]\left[\mathbf{h}(\boldsymbol{\chi}^{(i)-}_{k+1}) - \hat{\mathbf{z}}^{-}_{k+1}\right]^{\mathsf{T}} + \mathbf{R}_{k+1}\\ (\mathbf{P}^{-}_{\hat{\mathbf{x}}\hat{\mathbf{z}}})_{k+1} &= \frac{1}{M-1}\sum_{i=1}^{M}\left[\boldsymbol{\chi}^{(i)-}_{k+1} - \hat{\mathbf{x}}^{-}_{k+1}\right]\left[\mathbf{h}(\boldsymbol{\chi}^{(i)-}_{k+1}) - \hat{\mathbf{z}}^{-}_{k+1}\right]^{\mathsf{T}}.\end{aligned} \tag{18}$$

The Kalman gain is

$$\mathbf{K}_{k+1} = (\mathbf{P}^{-}_{\hat{\mathbf{x}}\hat{\mathbf{z}}})_{k+1}(\mathbf{P}^{-}_{\hat{\mathbf{z}}\hat{\mathbf{z}}})^{-1}_{k+1}. \tag{19}$$

For a given observation $\mathbf{z}_{k+1}$, the state update of ensemble member i is

$$\boldsymbol{\chi}^{(i)}_{k+1} = \boldsymbol{\chi}^{(i)-}_{k+1} + \mathbf{K}_{k+1}\left[\mathbf{z}_{k+1} + \boldsymbol{\eta}^{(i)}_{k+1} - \mathbf{h}(\boldsymbol{\chi}^{(i)-}_{k+1})\right], \tag{20}$$

where $\boldsymbol{\eta}^{(i)}_{k+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{k+1})$ models Gaussian measurement noise. Finally, the updated state estimate and the corresponding covariance matrix are

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\chi}^{(i)}_{k+1},\\ (\mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}})_{k+1} &= (\mathbf{P}^{-}_{\hat{\mathbf{x}}\hat{\mathbf{x}}})_{k+1} - \mathbf{K}_{k+1}(\mathbf{P}^{-}_{\hat{\mathbf{z}}\hat{\mathbf{z}}})_{k+1}\mathbf{K}^{\mathsf{T}}_{k+1}.\end{aligned} \tag{21}$$

During each update step, we assign the entry in each ensemble member that corresponds to the logarithm of the drug-caused mortality rate, $\tilde{\mu}_{\mathrm{d}} = \log(\mu_{\mathrm{d}})$, to be equal to the logarithm of the ratio of observed overdose fatalities and the ensemble mean of the population, while also accounting for the underlying noise.

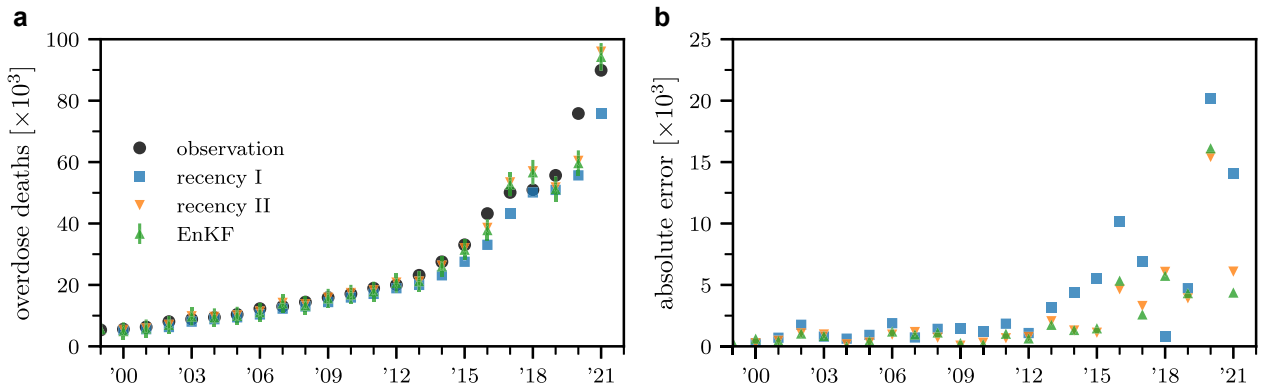

Fig. 9. Comparison of EnKF forecasts with forecasting heuristics. a) Nationwide overdose deaths in the United States as obtained from the CDC WONDER database (black disks) and corresponding forecasts (blue squares: recency heuristic I; orange inverted triangles: recency heuristic II; green triangles: EnKF) for the period 1999–2021. In recency heuristic I, the projected number of total overdose deaths in a specific age group in year Y is given by the number of overdose deaths in the same age group in year $Y-1$. In recency heuristic II, the projected number of overdose deaths in a specific age group in year Y is calculated as the sum of the number of overdose deaths in that age group in year $Y-1$ and the difference between the number of overdose deaths in the same age group in years $Y-1$ and $Y-2$. For each year, we calculate the total number of deaths by summing across all age groups. Error bars represent EnKF 3σ intervals. b) Absolute errors associated with both forecasting heuristics and the EnKF forecast.

Overdose mortality data

Overdose fatalities extracted from the CDC WONDER database were identified using the International Classification of Diseases, Tenth Revision (ICD-10) cause-of-death codes X40–44 (unintentional), X60–64 (suicide), X85 (homicide), Y10–14 (undetermined intent), and all other drug-induced causes. National-level and county-level data were extracted for the 1999–2020 period. The drug categories examined are poisoning by narcotics and psychodysleptics (hallucinogens) (T40) and by psychostimulants with abuse potential (T43.6). Specific subcategories analyzed within T40 are heroin (T40.1), natural and semisynthetic opioids (T40.2), and synthetic opioids other than methadone (T40.4). Deaths involving more than one drug type were included in each applicable category. Entries with an insufficient number of deaths were excluded.

Trends in population, overdose mortality, and drug types across three regions

As outlined in the main text, the evolution of overdose fatalities in Los Angeles County, Cook County, and the five boroughs of New York City exhibit distinct patterns over the past two decades. In Fig. 8, we provide additional insights into population, overdose mortality, and drug-type trends across these regions. Figure 8a–c shows the age-stratified population distribution in the three jurisdictions from 1999 to 2021. We observe that New York City exhibits a more pronounced peak of residents of ages between 20 and 30 years compared to the other two counties. Overdose deaths counts in Los Angeles County, shown in Fig. 8d, are characterized by a peak in the 20 to 40 age group that emerged over the past few years, aligning with the national trends shown in Fig. 1b. Conversely, Cook County and the five boroughs of New York City show the highest number of overdose deaths among older age groups as can be seen in Fig. 8e,f.

We also observe variations in the drug types that cause overdose deaths in the three regions. Figure 8g–i shows that crude rates for prescription opioids (T40.2) and heroin (T40.1) have remained relatively stable in Los Angeles County over the past two decades. However, significant increases in crude rates associated with overdoses due to synthetic opioids (T40.4, mainly fentanyl), and psychostimulants (T43.6, mainly methamphetamines) are observed over the last 5 years, indicating that these two drug classes have become the primary drivers of overdose mortality in this region. Fentanyl also emerges as a key driver of recent overdose mortality in Cook County and the five boroughs of New York City; contrary to Los Angeles County however, crude rates associated with heroin surpass those of both methamphetamines and prescription opioids.

In Los Angeles County, the crude rate for methamphetamines is large compared to that of the other regions throughout the two decades under investigation. This observation aligns with the fact that the first manufacturers and distributors of illegal methamphetamines in the United States were California-based biker gangs, particularly Hell's Angels, (98) whose activities began in the 1960s and remained mostly confined to the state for several decades. Also notice the temporary spike in fentanyl-driven overdoses in Cook County for 2005 and 2006. Indeed, Chicago was used as the first test market for fentanyl in the United States under the street names Drop Dead and Lethal Injection. Its 2005 introduction caused a temporary outbreak of deaths. Although authorities were able to stop distribution from the Mexico-based clandestine laboratory that produced the fentanyl responsible for these deaths in 2006, other organized criminal groups were able to resume the large-scale distribution of fentanyl across the country in later years (99, 100).

Comparison with baselines

We compare our EnKF forecasts with two baseline projections that are obtained by assuming a continuation of present trends. Such baseline projections have been shown to perform well especially over short forecasting time horizons (101, 102). We refer to these two baselines as "recency heuristics" (101, 103). In the first baseline (i.e. recency heuristic I), the forecast for the number of overdose deaths in a specific age group in year Y is assumed to be the same as the number of overdose deaths in the same age group in year $Y-1$. This heuristic is effective when overdose dynamics remain stable. In the second baseline (i.e. recency heuristic II), the forecast for the number of overdose deaths in a specific age group in year Y is calculated as the sum of the number of overdose deaths in that age group in year $Y-1$ and the difference between the number of overdose deaths in the same age group in years $Y-1$ and $Y-2$. This heuristic will perform better than recency heuristic I if the trend from the previous year persists and overdose dynamics are not stable.

Figure 9 shows the forecasts of nationwide overdose deaths upon combining all age groups for the different methods used and the corresponding forecasting errors. The years for which recency heuristic I performs better than the other two approaches (recency heuristics II and the EnKF) are those for which overdose counts do not change appreciably (e.g. 2018) compared to the previous year. Recency heuristic II is an effective forecasting method. However, it may exhibit a higher degree of overshooting compared to the EnKF forecast, as observed in 2021.

Since they rely on the most recent available observational data, model-free recency-type heuristics, such as recency heuristic I and II usually perform well when forecasting over short time horizons. Previous studies have demonstrated that recency heuristic I can provide more accurate forecasts than Google Flu Trends (101), and that recency heuristic II can compete favorably with CDC and ECDC COVID-19 ensemble forecasting models (102, 104). However, simple data-driven heuristics do not typically yield reliable long-term forecasts, as they lack the ability to account for underlying population-level dynamics. Unlike data-assimilation methods as the one used in our work, the described recency heuristics only offer point estimates without CIs. Additionally, as they are not based on an underlying mechanistic model, these heuristics are not suitable for inferring crucial epidemiological parameters like mortality and addiction rates.

Funding

T.C. and M.R.D. acknowledge financial support from the ARO through grant W911NF-18-1-0345. L.B. and M.R.D. acknowledge financial support from the ARO through grant W911NF-23-1-0129.

Author Contributions

L.B. conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, funding acquisition. T.C. conceptualization, methodology, formal analysis, writing—review and editing, visualization, funding acquisition. M.R.D. conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, funding acquisition.

Preprints

This manuscript was posted on a preprint server: https://doi.org/10.1101/2023.09.25.23296097.

Data Availability

Our source codes are publicly available at https://gitlab.com/ComputationalScience/overdose-da. The CDC WONDER portal can be accessed at https://wonder.cdc.gov/.

References

1 Mattson CL, *et al.* 2021. Trends and geographic patterns in drug and synthetic opioid overdose deaths—United States, 2013–2019. *Morb Mortal Wkly Rep*. 70(6):202–207.
2 O'Donnell J, Gladden RM, Mattson CL, Hunter C, Davis NL. 2020. Vital signs: characteristics of drug overdose deaths involving opioids and stimulants—24 states and the District of Columbia, January–June 2019. *Morb Mortal Wkly Rep*. 69(35): 1189–1197.
3 Jones CM, Einstein EB, Compton WM. 2018. Changes in synthetic opioid involvement in drug overdose deaths in the United States, 2010–2016. *J Am Med Assoc*. 319(17):1819–1821.
4 Armenian P, Vo KT, Barr-Walker J, Lynch KL. 2018. Fentanyl, fentanyl analogs and novel synthetic opioids: a comprehensive review. *Neuropharmacology*. 134:121–132.
5 Forman RF, Block LG. 2006. The marketing of opioid medications without prescription over the internet. *J Public Policy Mark*. 25(2):133–146.
6 Mackey TK, Kalyanam J, Katsuki T, Lanckriet G. 2017. Twitter-based detection of illegal online sale of prescription opioid. *Am J Public Health*. 107(12):1910–1915.
7 Lamy FR, *et al.* 2020. Listed for sale: analyzing data on fentanyl, fentanyl analogs and other novel synthetic opioids on one cryptomarket. *Drug Alcohol Depend*. 213:108115.
8 Duhart-Clarke SE, Kral AH, Zibbell JE. 2022. Consuming illicit opioids during a drug overdose epidemic: illicit fentanyls, drug discernment, and the radical transformation of the illicit opioid market. *Int J Drug Policy*. 99:103467.
9 Kariisa M, O'Donnell J, Kumar S, Mattson CL, Goldberger BA. 2023. Illicitly manufactured fentanyl involved overdose deaths with detected xylazine—United States, January 2019–June 2022. *Morb Mortal Wkly Rep*. 72(26):721–727.
10 Stein EM, Gennuso KP, Ugboaja DC, Remington PL. 2017. The epidemic of despair among White Americans: trends in the leading causes of premature death, 1999–2015. *Am J Public Health*. 107(10):1541–1547.
11 Case A, Deaton A. 2020. *Deaths of despair and the future of capitalism*. Princeton (NJ): Princeton University Press.
12 Friedman J, Akre S. 2021. COVID-19 and the drug overdose crisis: uncovering the deadliest months in the United States, January–July 2020. *Am J Public Health*. 111(7):1284–1291.
13 D'Orsogna MR, Böttcher L, Chou T. 2023. Fentanyl-driven acceleration of racial, gender and geographical disparities in drug overdose deaths in the United States. *PLoS Glob Public Health*. 3(3):e0000769.
14 Jalal H, *et al.* 2018. Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. *Science*. 361(6408):6408.
15 Peters DJ, Monnat SM, Hochstetler AL, Berg MT. 2020. The opioid hydra: understanding overdose mortality epidemics and syndemics across the rural–urban continuum. *Rural Sociol*. 85(3): 589–622.
16 Powell D, Shetty KD, Peet ED. 2023. Trends in overdose deaths involving gabapentinoids and Z-drugs in the United States. *Drug Alcohol Depend*. 249:109952.
17 Segel JE, Winkelman TNA. 2021. Persistence and pervasiveness: early wave opioid overdose death rates associated with subsequent overdose death rates. *Public Health Rep*. 136(2):212–218.
18 Blanco C, Wall MM, Olfson M. 2022. Data needs and models for the opioid epidemic. *Mol Psychiatry*. 27(2):787–792.
19 Lim TY, *et al.* 2022. Modeling the evolution of the US opioid crisis for national policy development. *Proc Natl Acad Sci U S A*. 119(23): e2115714119.
20 Borquez A, Martin NK. 2022. Fatal overdose: predicting to prevent. *Int J Drug Policy*. 104:103677.
21 Monnat SM. 2018. Factors associated with county-level differences in US drug-related mortality rates. *Am J Prev Med*. 54(5): 611–619.
22 Rigg K, Monnat SM, Chavez MN. 2018. Opioid-related mortality in rural America: geographic heterogeneity and intervention strategies. *Int J Drug Policy*. 57:119–129.

23 Brownstein JS, Green TC, Cassidy TA, Butler SF. 2010. Geographic information systems and pharmacoepidemiology: using spatial cluster detection to monitor local patterns of prescription opioid abuse. *Pharmacoepidemiol Drug Saf*. 19(6): 627–637.

24 Basak A, Cadena J, Marathe A, Vullikanti A. 2019. Detection of spatiotemporal prescription opioid hot spots with network scan statistics: multistate analysis. *JMIR Public Health Surveill*. 5(2):e12110.

25 Campo DS, Gussler JW, Sue A, Skums P, Khudyakov Y. 2020. Accurate spatiotemporal mapping of drug overdose deaths by machine learning of drug-related web-searches. *PLoS One*. 15(12):e0243622.

26 Marks C, *et al*. 2021. Identifying counties at risk of high overdose mortality burden during the emerging fentanyl epidemic in the USA: a predictive statistical modelling study. *Lancet Public Health*. 6(10):e720–e728.

27 Marks C, *et al*. 2021. Methodological approaches for the prediction of opioid use-related epidemics in the United States: a narrative review and cross-disciplinary call to action. *Transl Res*. 234:88–113.

28 Sumetsky N, *et al*. 2021. Predicting the future course of opioid overdose mortality: an example from two US states. *Epidemiology*. 32(1):61–69.

29 Wagner NM, *et al*. 2021. Development and validation of a prediction model for opioid use disorder among youth. *Drug Alcohol Depend*. 227:108980.

30 Stringfellow EJ, *et al*. 2022. Reducing opioid use disorder and overdose deaths in the united states: a dynamic modeling analysis. *Sci Adv*. 8(25):eabm8147.

31 Crassidis JL, Junkins JL. 2004. *Optimal estimation of dynamic systems*. Boca Raton (FL): Chapman and Hall/CRC.

32 Law K, Stuart A, Zygalakis K. 2015. *Data assimilation: a mathematical introduction*. Texts in Applied Mathematics. Cham, Switzerland: Springer.

33 Böttcher L, Chou T, D'Orsogna MR. 2023. Modeling and forecasting age-specific overdose mortality in the United States. *Eur Phys J Spec Top*. 232:1743.

34 Bauer P, Thorpe A, Brunet G. 2015. The quiet revolution of numerical weather prediction. *Nature*. 525(7567):47–55.

35 Lillacci G, Khammash M, Asthagiri AR. 2010. Parameter estimation and model selection in computational biology. *PLoS Comput Biol*. 6(3):e1000696.

36 Schneider T, *et al*. 2022. Epidemic management and control through risk-dependent individual contact interventions. *PLoS Comput Biol*. 18(6):e1010171.

37 Pei S, Liljeros F, Shaman J. 2021. Identifying asymptomatic spreaders of antimicrobial-resistant pathogens in hospital settings. *Proc Natl Acad Sci U S A*. 118(37):e2111190118.

38 Li R, *et al*. 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*. 368(6490):489–493.

39 Bomfim R, *et al*. 2020. Predicting dengue outbreaks at neighbourhood level using human mobility in urban areas. *J R Soc Interface*. 17(171):20200691.

40 M'Kendrick AG. 1925. Applications of mathematics to medical problems. *Proc Edinburgh Math Soc*. 44:98–130.

41 Kermack WO, McKendrick AG. 1991. Contributions to the mathematical theory of epidemics—I. 1927. *Bull Math Biol*. 53:33–55.

42 Kermack WO, McKendrick AG. 1991. Contributions to the mathematical theory of epidemics—II. The problem of endemicity. *Bull Math Biol*. 53:57–87.

43 Kermack WO, McKendrick AG. 1991. Contributions to the mathematical theory of epidemics—III. Further studies of the problem of endemicity. *Bull Math Biol*. 53:89–118.

44 Diekmann O, Othmer HG, Planqué R, Bootsma MCJ. 2021. The discrete-time Kermack–McKendrick model: a versatile and computationally attractive framework for modeling epidemics. *Proc Natl Acad Sci U S A*. 118(39):e2106332118.

45 Xia M, Greenman C, Chou T. 2020. PDE models of adder mechanisms in cellular proliferation. *SIAM J Appl Math*. 80(3): 1307–1335.

46 Wang Y, Dessalles R, Chou T. 2022. Modeling the impact of birth control policies on China's population and age: effects of delayed births and minimum birth age constraints. *R Soc Open Sci*. 9(6):211619.

47 Schenzle D. 1984. An age-structured model of pre- and post-vaccination measles transmission. *Math Med Biol*. 1(2):169–191.

48 Castillo-Chavez C, Feng Z. 1998. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. *Math Biosci*. 151(2):135–154.

49 Rong L, Feng Z, Perelson AS. 2007. Mathematical analysis of age-structured HIV-1 dynamics with combination antiretroviral therapy. *SIAM J Appl Math*. 67(3):731–756.

50 Böttcher L, Xia M, Chou T. 2020. Why case fatality ratios can be misleading: individual-and population-based mortality estimates and factors influencing them. *Phys Biol*. 17(6):065003.

51 Chuang YL, Chou T, D'Orsogna MR. 2018. Age-structured social interactions enhance radicalization. *J Math Sociol*. 42(3):128–151.

52 Yang J, Li X, Zhang F. 2016. Global dynamics of a heroin epidemic model with age structure and nonlinear incidence. *Int J Biomath*. 9:1650033.

53 Liu L, Liu X. 2019. Mathematical analysis for an age-structured heroin epidemic model. *Acta Appl Math*. 164(1):193–217.

54 Chekroun A, Frioui MN, Kuniya T, Touaoula TM. 2020. Mathematical analysis of an age structured heroin–cocaine epidemic model. *Discrete Continuous Dyn Syst Ser B*. 25:4449–4477.

55 Din A, Li Y. 2020. Controlling heroin addiction via age-structured modeling. *Adv Differ Equ*. 2020:521.

56 Duan X, Cheng H, Martcheva M, Yuan S. 2021. Dynamics of an age structured heroin transmission model with imperfect vaccination. *Int J Bifurcat Chaos*. 31(10):2150157.

57 Khan A, Zaman G, Ullah R, Naveed N. 2021. Optimal control strategies for a heroin epidemic model with age-dependent susceptibility and recovery-age. *AIMS Math*. 6(2):1377–1394.

58 Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res Oceans*. 99(C5):10143–10162.

59 Gomes T, *et al*. 2022. *Patterns of medication and healthcare use among people who died of an opioid-related toxicity during the COVID-19 pandemic in Ontario*. Toronto (ON): The Ontario Drug Policy Research Network.

60 Palis H, *et al*. 2022. Association of opioid and stimulant use disorder diagnoses with fatal and nonfatal overdose among people with a history of incarceration. *JAMA Netw Open*. 5(11): e2243653.

61 Gompertz B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc Lond*. 115: 513–583.

62 Makeham WM. 1860. On the law of mortality and the construction of annuity tables. *J Inst Actuar*. 8(6):301–310.

63 Siler W. 1979. A competing-risk model for animal mortality. *Ecology*. 60(4):750–757.

64 Siler W. 1983. Parameters of mortality in human populations with widely varying life spans. *Stat Med*. 2(3):373–380.
65 Cohen JE, Bohk-Ewald C, Rau R. 2018. Gompertz, Makeham, and Siler models explain Taylor's law in human mortality data. *Demogr Res*. 38:773–842.
66 Substance Abuse and Mental Health Services Administration. 2011. Results from the 2010 National Survey on Drug Use and Health: summary of national findings. NSDUH Series H-41, HHS Publication No. (SMA) 11-4658. Rockville (MD): Substance Abuse and Mental Health Services Administration.
67 Mathers BM, *et al.* 2013. Mortality among people who inject drugs: a systematic review and meta-analysis. *Bull World Health Organ*. 91(2):102–123.
68 Lindblad R, *et al.* 2016. Mortality rates among substance use disorder participants in clinical trials: pooled analysis of twenty-two clinical trials within the national drug abuse treatment clinical trials network. *J Subst Abuse Treat*. 70:73–80.
69 Hser YI, *et al.* 2017. High mortality among patients with opioid use disorder in a large healthcare system. *J Addict Med*. 11(4): 315–319.
70 Sordo L, *et al.* 2017. Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies. *Br Med J*. 357:j1550.
71 King C, Cook R, Korthuis PT, Morris C, Englander H. 2022. Causes of death in the 12 months after hospital discharge among patients with opioid use disorder. *J Addict Med*. 16(4):466–469.
72 Bahji A, Cheng B, Gray S, Stuart H. 2020. Mortality among people with opioid use disorder: a systematic review and meta-analysis. *J Addict Med*. 14(4):e118–e132.
73 Jalal H, Burke DS. 2021. Carfentanil and the rise and fall of overdose deaths in the United States. *Addiction*. 116(6):1593–1599.
74 Visconti AJ, Santos G-M, Lemos NP, Burke C, Coffin PO. 2015. Opioid overdose deaths in the city and county of San Francisco: prevalence, distribution, and disparities. *J Urban Health*. 92(4):758–772.
75 Marshall JR, *et al.* 2019. Socioeconomic and geographical disparities in prescription and illicit opioid-related overdose deaths in Orange County, California, from 2010–2014. *Subst Abus*. 40(1): 80–86.
76 Romeiser JL, Labriola J, Meliker JR. 2019. Geographic patterns of prescription opioids and opioid overdose deaths in New York State, 2013–2015. *Drug Alcohol Depend*. 195:94–100.
77 Marotta PL, *et al.* 2019. Assessing spatial relationships between prescription drugs, race, and overdose in New York State from 2013 to 2015. *J Psychoactive Drugs*. 51(4):360–370.
78 Kline D, Pan Y, Hepler SA. 2021. Spatiotemporal trends in opioid overdose deaths by race for counties in Ohio. *Epidemiology*. 32(2): 295–302.
79 King A, *et al.* 2019. Carfentanil-associated mortality in Wayne County, Michigan, 2015–2017. *Am J Public Health*. 109(2):300–302.
80 Lister JJ, Ellis JD, Yoon M. 2020. Opioid prescribing and opioid-overdose deaths in Michigan: urban-rural comparisons and changes across 2013–2017. *Addict Behav Rep*. 11:100234.
81 Rossen LM, Bastian B, Warner M, Khan D. 2023. NCHS—Drug Poisoning Mortality by County: United States; 1999–2018 [accessed 2024 Feb 22]. https://data.cdc.gov/NCHS/NCHS-Drug-Poisoning-Mortality-by-County-United-Sta/rpvx-m2md.
82 Lister JJ, Weaver A, Ellis JD, Himle JA, Ledgerwood DM. 2020. A systematic review of rural-specific barriers to medication treatment for opioid use disorder in the United States. *Am J Drug Alcohol Abuse*. 46(3):273–288.
83 Post LA, *et al.* 2022. Geographic trends in opioid overdoses in the US from 1999 to 2020. *JAMA Netw Open*. 5(7):e2223631.
84 Gini C. 1912. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. Bologna (Italy): C. Cuppini.
85 Xu S, Böttcher L, Chou T. 2020. Diversity in biology: definitions, quantification and models. *Phys Biol*. 17(3):031001.
86 Courtwright DT. 2001. *Dark paradise: a history of opiate addiction in America*. Cambridge (MA): Harvard University Press.
87 Musto DF. 1999. *The American disease: origins of narcotic control*. 3rd ed. Oxford (UK): Oxford University Press.
88 Courtwright DT. 1992. A century of American narcotic policy. In: Gerstein DR, Harwood HJ, editors. *Treating drug problems*. Vol. 2. Washington (DC): National Academies Press.
89 Jones CM, Warner M, Hedegaard H, Compton W. 2019. Data quality considerations when using county-level opioid overdose death rates to inform policy and practice. *Drug Alcohol Depend*. 204:107549.
90 Chou T, Greenman CD. 2016. A hierarchical kinetic theory of birth, death and fission in age-structured interacting populations. *J Stat Phys*. 164(1):49–76.
91 Anderson JL. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon Weather Rev*. 129:2884–2903.
92 Evensen G, Van Leeuwen PJ. 2000. An ensemble Kalman smoother for nonlinear dynamics. *Mon Weather Rev*. 128: 1852–1867.
93 National Center for Health Statistics. 2021. *NCHS Fact Sheet—June 2021: NCHS data on drug overdose deaths*. Report. National Center for Health Statistics.
94 Lipari RN, Park-Lee E, Substance Abuse and Mental Health Services Administration. 2019. Key substance use and mental health indicators in the United States: results from the 2018 National Survey on Drug Use and Health (HHS Publication No. PEP19-5068, NSDUH Series H-54). Rockville (MD): Center for Behavioral Health Statistics and Quality.
95 Brown RG, Hwang PYC. 1997. *Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions*. Hoboken (NJ): Wiley.
96 Labbe R. 2022. Kalman and Bayesian Filters in Python [accessed 2024 Feb 22]. https://github.com/rlabbe/Kalman-and-Bayesian-Filters-in-Python/blob/master/Appendix-E-Ensemble-Kalman-Filters.ipynb.
97 Brown RG, Hwang PYC. 2012. *Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions*. Hoboken (NJ): Wiley.
98 Quinones S. 2015. *Dreamland: the true tales of America's opiate epidemic*. New York (NY): Bloomsbury Publishing.
99 Quinones S. 2021. *The least of us: true tales of America and hope in the time of fentanyl and meth*. New York (NY): Bloomsbury Publishing.
100 Quinones S. 2023. *America and hope in the time of fentanyl and meth: a fireside chat with Sam Quinones*. Syracuse (NY): Center for Policy Research.
101 Katsikopoulos KV, Şimşek Ö, Buckmann M, Gigerenzer G. 2022. Transparent modeling of influenza incidence: big data or a single data point from psychological theory? *Int J Forecast*. 38(2): 613–619.
102 Antulov-Fantulin N, Böttcher L. 2022. On the accuracy of short-term COVID-19 fatality forecasts. *BMC Infect Dis*. 22(1):1–7.
103 Gigerenzer G, Todd PM, the ABC Research Group. 1999. *Simple heuristics that make us smart*. New York (NY): Oxford University Press.
104 Chharia A, *et al.* 2022. Accuracy of US CDC COVID-19 forecasting models. *medRxiv* 2022.04.20.22274097. https://doi.org/10.1101/2022.04.20.22274097.