Editorial

Fentanyl in the US heroin supply: A rapidly changing risk environment

The supply of heroin into the US has changed with new source-forms and market strategies. Of particular concern is the rampant and persistent adulteration of heroin with synthetic opioids, most conspicuously the family of fentanyl. Medical consequences including alarming reports of increasing opioid-related overdose, chest wall rigidity syndrome and naloxone resistant overdoses demand urgent public health and policy responses. This Special Section of the International Journal of Drug Policy explores the phenomenon of substituting “traditional” heroin with new less predictable and less culturally established forms and how those forms are perceived and adopted or avoided by those who use them; it maps the consequences and explores structural risk as well as behavioral risk-factors. The research papers and commentaries presented herein discuss the supply-side shock of synthetic opioids that presents multiple overlapping challenges and paradoxes for epidemiological and toxicological surveillance, supply-side intervention, cryptomarkets and public health intervention.

A crisis of historic scale

Deaths due to illicit drug overdose in the United States have reached historic proportions with annual numbers exceeding those due to motor vehicle accidents, gun violence or HIV/AIDS — at the height of the US HIV crisis in the early 1990s (Katz, 2017a, 2017b). This is driving up the US mortality rate (2015) for the first time since 1999 (Xu, Murphy, Kochanek, & Arias, 2016). The most common cause of drug overdose deaths is opioids with 33,000 opioid deaths in 2015 (Centers for Disease Control and Prevention, 2016). This is a triple epidemic with rising waves of deaths due to separate types of opioids each building on top of the prior wave. The first wave of prescription opioid mortality began in the 1990s. The second wave, due to heroin, began around 2010 with heroin-related overdose deaths tripling since then (Warner, Trinidad, Bastian, Miniño, & Hedegaard, 2016). Now synthetic opioid-related overdoses, including those due to illicitly manufactured fentanyl and fentanyl analogues, are causing the third wave with these overdose deaths doubling between 2013 and 2014 (Frank & Pollack, 2017; Gladden, Martinez, & Seth, 2016; Warner et al., 2016).

Heroin- and fentanyl-related overdose are also rising worldwide. Drug overdose deaths are up 327% in Canada, 64% in the United Kingdom and 61% in Australia (Harm Reduction International, 2017) although Australia’s fentanyl-related overdose deaths are due to illicit use of pharmaceutical fentanyl not illicitly manufactured fentanyl (Latimer, Ling, Flaherty, Jauncey, & Salmon, 2016; Roxburgh et al., 2013). There have been 1,019 fentanyl-detected deaths in Canada as of 2014 with more than half occurring 2013–2014 (Canadian Community Epidemiology Network on Drug Use, 2015). This Special Section focuses solely on the US situation and only on the overdose problem as it relates to heroin and fentanyl. Other drug injection injuries and infections, e.g. SSTI, endocarditis and hepatitis C, are also on the rise in the US (Ciccarone, Unick, Cohen, Mars, & Rosenblum, 2016; Larney, Peacock, Mathers, Hickman, & Degenhardt, 2017; Zibbell et al., 2015; Ciccarone et al., 2016).

The origin story of this triple or intertwined epidemic has been summed up in a number of papers (Compton, Jones, & Baldwin, 2016; Kolodny et al., 2015; Madras, 2017; Mars, Bourgois, Karon, Montero, & Ciccarone, 2014; Unick, Rosenblum, Mars, & Ciccarone, 2013) which largely focus on excessive physician prescribing practices. However there are deeper structural roots to the opioid crisis, e.g. economic distress (Monnat, 2016), that require further exploration. The role of positive supply shocks, both licit and illicit, in causing drug use epidemics, is well recognized. In the US when morphine and heroin were introduced in the late 19th Century as legitimate medical products, widespread addiction followed (Courtwright 2001). In the 1970s a US inner-city heroin epidemic was aided by the introduction of a highly refined heroin from South East Asia (McCoy, 2003, revised edition). Likewise in the 1990s a new source-form of heroin, this time from Colombia, caused a significant decline in heroin price and a wave of heroin consequences (Ciccarone, Unick, & Kraus, 2009).

There are three intersecting possible drivers for rising heroin-related overdose in the US: rising numbers of users, a heroin positive supply shock and increasing dangers due to fentanyl in the heroin supply. The number of heroin users as reported in a national survey has doubled since 2008 (Center for Behavioral Health Statistics and Quality, 2015). Increased availability of heroin is not to be underestimated: reports from the US Drug Enforcement Administration highlight an increase in heroin seizures, 143% from 2010 to 2015, while heroin price per mg pure has declined to historically low levels (Drug Enforcement Administration, 2016a). Importantly, they reveal structural market changes that represent an under-recognized heroin supply shock. Looking at wholesale data, in 2014 an impressive 79% of US heroin was Mexican sourced — up from 15% a decade prior (Drug Enforcement Administration, 2016b). Strikingly, there is a new form of heroin from Mexico that is a mimic of South American heroin; a refined powder heroin quite dissimilar to the crude heroin that Mexican drug cartels sold for decades in the US. This new source-form of heroin is of high quality.

http://dx.doi.org/10.1016/j.drugpo.2017.06.010
0955-3959/© 2017 Elsevier B.V. All rights reserved.
with an average wholesale purity of 74%; higher than South American and much higher than traditional Mexican crude heroin. That this new heroin mimic is heading to the eastern US is another revelation: it breaks the decades old division of the US heroin market with Colombian-sourced heroin heading east and Mexican-sourced heroin going west (Ciccarone, 2009; Ciccarone et al., 2009). According to the DEA, Mexican drug trafficking organizations are “aggressively expanding and taking increasing control of eastern US heroin markets” (Drug Enforcement Administration, 2016b). In addition, analyses of retail level heroin samples show a significant number of samples of “unknown” origin; most of these samples are found in eastern US cities. The average retail price per mg-pure of heroin has almost halved 2010–2014 (Drug Enforcement Administration, 2016b). The disruption in the US heroin market also includes new routing and marketing strategies that facilitate heroin getting to towns and rural locations outside the typical inner-city domain (Quinones, 2015).

A second positive supply side shock, beginning in 2013, comes in the wave of synthetic opioids contaminating or circulating within the heroin supply. According to drug seizure data from the DEA National Forensics Laboratory Information Service, fentanyl reports increased 1400% from 2013 (934) to 2015 (14,051) (Drug Enforcement Administration Diversion Control Division, 2016). Fentanyl was the number one synthetic opioid reported in 2015 with acetyl fentanyl second (2142). The family of fentanyl has a range of potency from roughly 3–10,000 times that of morphine by weight (Suzuki & El-Haddad, 2017): given that fentanyl and heroin are approximately 100 and 2.5 times the potency of morphine, respectively, fentanyl is approximately 40 times stronger than heroin per mg. Volume of seized fentanyl highlights the magnitude of the structural risk increase. According to the Office of National Drug Control Policy (ONDCP) 668 kg of fentanyl powder, originating from China, were seized in 2016 (preliminary data), representing a 426% increase from the prior year (Baum, 2017). Even given conservative estimates of 5% purity and a high seizure to importation ratio of 1:4 (Paoli, Greenfield, & Reuter, 2009), at least 134 kg of pure fentanyl entered the US in 2016. With an estimate of 0.25–10 mg as a replacement dose for a single dose of heroin, this represents 134–536 million replacement doses – about one a day for the estimated million heroin users in the US (United Nations Office of Drug Control, 2016). As supply has increased, the number of fentanyl-related overdose deaths has doubled from 2012 to 2014 (Warner et al., 2016); with a strong correlation between synthetic opioid overdose deaths and fentanyl seizure reports in 27 US states with available data (Gladden et al., 2016).

Contributions to this special section

Taken together the above market factors represent a positive supply shock to the US heroin market as well as a significant shift in the structural risk environment for those who use heroin. The US Centers for Disease Control and Prevention (CDC) identified eight high-burden states in which the crude death rate due to synthetic opioids increased 174% from 2013 to 2014. These eight states were located in the US regions of the Northeast, Midwest and South. The research papers in the Special Section highlight this regional pattern in the intertwined heroin and fentanyl epidemics.

Unick and Ciccarone (2017), utilizing national hospitalization data, make geographic and demographic distinctions between heroin- and prescription opioid-related overdose. Their analyses reveal the strong regionality of the heroin overdose epidemic: the US Census divisions identified in their data as having the greatest increase in heroin overdose contain those states identified by the CDC as having the highest increase in rates of synthetic opioid overdose. This suggests the intertwining of heroin and fentanyl supply, use and consequences at a US regional level (Unick & Ciccarone, 2017). These authors also point out that heroin overdose is skewing to younger age groups than with opioid pill overdose and becoming more ethnically diverse in certain regions.

Slavova et al. (2017) reporting on data from the “high-burden” state of Kentucky also find an age shift for fentanyl- and heroin-related overdose to younger age groups over time. Fentanyl- and heroin-related overdose deaths increased 2011–2015 with the highest rates among 25–34 year-olds. Examining emergency department data, the number of rural heroin overdose visits exceeded those for opioid pill overdose for the first time in mid-2016. The authors report a strong positive correlation between crime lab positive fentanyl submissions and the number of fentanyl-related overdose deaths at the state level. Untested by the authors, but visually striking (to this Editor), is the possible correlation between fentanyl submissions and heroin-related overdose emergency visits (Slavova et al., 2017, Fig. 3) showing the fungibility in this supply and consequence conundrum. Slavova et al. (2017) report another compelling finding: the contrast between the ratio for emergency visits and deaths between fentanyl-related overdose (1:1 ratio) and heroin-related overdose (10:1). The authors report the likely explanation of emergency department under-reporting of fentanyl visits. The data also support another possibility: a narrowed therapeutic window for fentanyl overdose response (see interventions below).

In Rhode Island, a state in the US region of New England, the number and proportion of drug overdose deaths involving fentanyl has increased sharply between 2014 and 2016. Marshall et al. (2017) report that fentanyl is now the primary or contributing cause of death in over half of all overdose fatalities, with 28% of decedents aged 18–29. Their geospatial analyses demonstrate a widespread distribution of fentanyl-related overdose suggesting a broad (and top-down) supply distribution.

Two papers emphasize the need to understand the perspectives of persons who use heroin on use of and adaptation to fentanyl and fentanyl-adulterated heroin. Carroll, Marshall, Rich, and Green (2017) report on data from an interdisciplinary study of persons using illicit opioids in Rhode Island. In their survey study, participants reported that exposure to fentanyl was common and that the biggest independent risk for suspected fentanyl exposure was regular use of heroin. Heroin users often suspect they have been exposed to fentanyl but are unsure. While there may be uncertainty on both sides of the dealing transaction regarding the presence of fentanyl in the heroin sold, users try to minimize this uncertainty through established relationships with their dealers.

Fentanyl adulteration of heroin represents a significant change in the structural risk environment for users of heroin (Ciccarone, Ondocin, & Mars, 2017). Ciccarone and colleagues report on qualitative findings from Massachusetts, a CDC “high-burden” state for synthetic overdose deaths, and point out their respondents’ discernment of and mixed opinions regarding the desirability for fentanyl. This contrasts somewhat with Carroll et al.’s report on their respondents’ limited ability to discern fentanyl and low desirability for it (Carroll et al., 2017). Clearly use patterns and perceptions are moving targets with local and evolving cultural norms and degrees of adaptation. Pharmaceutical fentanyl has been reported to have some desirability in survey studies (Firestone, Goldman, & Fischer, 2009) as well as controlled settings (Greenwald, 2008), but in the current epidemic the fentanyl is illicitly manufactured (Sommerville et al., 2017) and surreptitiously sold, not self-chosen pharmaceutical product.

Ciccarone et al.’s (2017) participants report increased availability of heroin and changes in heroin characteristics and form since 2013. The authors identify a possible heroin and fentanyl typology based on participants’ responses and call for its empiric testing. The authors also describe users’ methods of identifying fentanyl in
their heroin, with an emergent hierarchy of reliability: embodied effects were considered the most reliable, followed by taste (on the tongue), solution appearance and finally, powder appearance.

Four commentaries received for this Special Section discuss the policy and interventional challenges and opportunities in addressing the rising risk environment imposed by fentanyl. Beletsky and Davis (2017) detail the unintended consequences of supply side interventions on the chain of events connecting the opioid pill to heroin to fentanyl epidemics. They argue that the emergence of illicit synthetic opioids is the result of the “Iron Law of Prohibition” in which the potency by weight/volume of a drug or alcohol product goes up as economic pressure from interdiction forces volume of the supply flow to shrink. In the case of this intertwined epidemic, prescribing restrictions forced a majority of dependent users to more potent and available street heroin and in the larger interdiction sphere, supply pressure on heroin is forcing a market change to more potent synthetics with routing options, e.g. cryptomarkets and postal delivery, less available to heroin. This “law” also supports a point made earlier: Mexican-sourced heroin is shifting from a bulkier crude product, i.e., ‘black tar’ heroin to a refined powder possibly due to the same paradoxical pressures.

The role of cryptomarkets in fentanyl supply is elucidated by Gilbert and Dasgupta (2017) who remind us that fentanyl outbreaks have occurred previously in the US and Europe, as early as the 1990s. The challenge with fentanyl and other synthetics is compounded by new sources and poorly identified and understood supply chains. Cryptomarkets combine communication protocols and “block-chain currencies” to deliberately obfuscate both suppliers’ and buyers’ geographic and personal identities while facilitating transactions of money and product. For obvious reasons these markets are poorly understood and yet are in competition with cartel-based drug markets. Examples of cryptomarkets given by the authors highlight the diversity of products and routing. The authors also highlight the role of online community-based ethics and possibilities for harm reduction innovation imbedded in these disruptive technological endeavors.

Lucyk and Nelson (2017) discuss the challenges for toxicological surveillance of the evolving palette of synthetic opioids in the marketplace. For instance, states without centralized medical examiner systems specified the specific drug involved in a death much less frequently than states with centralized systems. The challenges of surveillance in emergency department settings, also emphasized by Slavova et al. (2017) in this issue, are compounded by logistics, costs and clinical need (Lucyk & Nelson, 2017). The authors recommend that clinicians take greater account of opioid users’ novel opioid experiences to initiate toxicity testing. The authors recommend adherence to guidelines regarding investigating deaths, use of more advanced analytic testing and non-targeted screening protocols.

Interventions to address the fentanyl–meets-heroin epidemic need creative expansion. Fairbairn, Coffin, and Walley (2017) highlight the evolution of community based overdose prevention programs, including peer-distribution of naloxone, as well as the present challenges and innovations required to move forward. The challenges in the synthetic era include erratic drug potency, naloxone dosing quandaries and a narrower overdose response window. Ciccarone et al. (2017) also discuss the overdose risk due to vicissitudes in drug potency, now magnified when potent opioids are combined. Fairbairn et al. (2017) call for scaling up naloxone distribution into new venues and improving post-overdose care, drug checking services and advanced opioid substitution including heroin prescription.

In addition to overdose education and naloxone distribution programs, a number of harm reduction interventions are discussed by the collective of authors in this Special Section. Carroll et al. (2017) and Ciccarone et al. (2017) discuss the effect that controlled consumption of drug, aka a “tester shot”, “taste” or insufflation (aka snorting) prior to or in lieu of injection might have in reducing overdose risk. Point-of-use testing or “drug checking” has been used especially in Europe, to inform and engage persons who consume drugs as well as provide a potential surveillance platform (Brunt et al. 2017; Caudevilla et al. 2016; Hondebrink, Nugteren-van Lonkhuyzen, Van Der Gouwe, Brunt, 2015). Rapid testing exists for some of the fentanyl and several contributors point out the positives and the challenges of this harm reduction strategy (Ciccarone et al., 2017; Fairbairn et al., 2017; Gilbert & Dasgupta, 2017).

Supervised injection facilities have a strong evidence base regarding reduction in harms and promotion of substance treatment and clinical services (Potier, Laprèvote, Dubois-Arber, Cottencin, & Rolland, 2014). The InSite program in Vancouver, Canada was successful in reducing population level fatal overdose rates (Marshall, Milloy, Wood, Montaner, & Kerr, 2011). Given that fentanyl overdose is a rapid event with a narrow therapeutic window (Green & Gilbert, 2016; Somerville et al., 2017) supervised injection facilities make sense in addressing the fentanyl overdose epidemic, yet they are underutilized in the hot spot of the epidemic — North America. In addition, innovations in supervised injection facilities are needed, e.g., in location, size, adaptability, cost, mobility etc. These programs are expanding in innovative ways due to the fentanyl overdose epidemic in Canada and some locations in the US are considering them.

Conclusion

Heroin source and form are part of the structural risk environment for users (Ciccarone, 2005, 2009), which has been worrissomely transformed by the phenomenon of fentanyl-adulterated -substituted heroin (FASH) (Ciccarone et al., 2017). An understanding of the heroin supply, and how it is changing, supports the geographic correlation between rising heroin overdose and availability of fentanyls. A ready explanation is that illicit fentanyls come in powder form and thus more easily adulterate powder heroin than solid, i.e., ‘black tar’, heroin (Carroll et al. 2017); the former predominates in the eastern US while the latter is more common in the West (Ciccarone, 2009). A new source-form of heroin from Mexico, supplying the eastern US, is an unknown entity requiring further inquiry. The “Iron Law of Prohibition”, as well described by Beletsky and Davis (2017), is likely in play as Mexican-sourced opioids transform from crude solid to refined powder heroin and then to fentanyl, along the way increasing in potency and risk.

Surveillance is a foundational aspect of public health. Current surveillance in this epidemic is simply not timely enough to inform rapid public health responses. Standardization of testing and reporting procedures, especially in emergency settings, is necessary along with development of syndromic systems and statistical models to achieve near-real time data monitoring (Lucyk & Nelson, 2017; Slavova et al., 2017). Fentanyl contamination of heroin can be seen as an evolving poisoning epidemic. Real time drug surveillance can provide timely data to inform public health and public safety interventions especially as substances change rapidly (Harris, Forseth, & Rhodes, 2015; Slavova et al., 2017). People who use heroin as well as front-line service providers have a need to know what is in the drug supply. In addition, the role of point-of-use drug testing in the fentanyl outbreak needs to be investigated.

Fentanyl and other synthetic opioids will likely become continuous or recurrent aspects of the heroin supply; indeed they may become the “new norm.” Analogous to the situation of levamisole in the cocaine supply: a ubiquitous adulterant which boosts the stimulant effect and which is now a normalized aspect
of illicit cocaine. Civil and economic development in source countries may lead to further decline in the illicit production of opium. Whereas the US once had four sources of heroin competing for retail market share, it is now down to two, with the oligopolistic trade of Mexican- and Colombian-sourced heroin potentially transforming to a monopolistic trade of opioids – including fentanyl – sourced solely from Mexico. In addition, the forces of globalization and neoliberal economic policies lead to increased flows of licit and illicit goods inhibiting drug interdiction at the borders (Ciccarone, 2005, 2009).

The crisis that is occurring requires a vital response, such as outlined by the contributors to this Special Section, including creative expansion of harm reduction strategies and technologies, expansion of evidence-based substance use treatment and improved public health surveillance. Overdose reduction policies will have their greatest impact if they take into account the perceptions and behaviors of the users themselves and promote tailored interventions that initiate indigenously and which engage users intuitively.

Acknowledgements

I am indebted to Sarah Mars for contributions to and fact checking and editing of this manuscript. I thank the entire Heroin in Transition research team for their dedicated work: Georgiy Bobashev, Philippe Bourgois, Mary Howie, Sarah Mars, Fernando Montero, Jeff Onodcin, Dan Rosenblum, George Jay Unick and Eliza Wheeler. A special thanks to Jon E. Zibbell for early and crucial guidance in the development of this Special Section. And thanks to Tim Rhodes at JUDP for planting the seed for this project and to Dan Norman at Elsevier for superb management of it especially in the final stages. Finally, I gratefully acknowledge research funding from US NIH/NIDA, Grant DA037820 (PI: Ciccarone).

References


Daniel Ciccarone*
University of California, San Francisco, Department of Family and Community Medicine, MU-3E, Box 900, 500 Parnassus Ave., San Francisco, CA 94143, United States

* Fax: +1 415 476 6051.
E-mail address: ciccaron@fcm.ucsf.edu (D. Ciccarone).