Please cite the Published Version

Additional Information: This is an Open Access article published in Policing: A Journal of Policy and Practice.

Policing and Mental ill-health: Using Big Data to Assess the Scale and Severity of, and the Frontline Resources Committed to, mental ill-health-related calls-for-service

Samuel Langton ⓘ *,**, Jon Bannister*, Mark Ellison*, Muhammad Salman Haleem*** and Karolina Krzemieniewska-Nandwani*

Abstract   Addressing public safety and welfare, inclusive of responding to incidents involving persons with mental ill-health (PMIH) has become an integral dimension of, and a significant challenge to, contemporary policing. Yet, little is known of the scale and severity of such PMIH-related policing demand, nor of the extent of frontline resource consumed in resolving such incidents. To address this shortfall, we deploy a bespoke text mining algorithm on police incident logs to estimate the proportion and severity of calls-for-service involving PMIH in a study of Greater Manchester, UK. Furthermore, and using Global Positioning System data, we then assess the amount of time spent by frontline officers responding to these calls. Findings suggest that existing police recording practices serve to significantly underestimate the scale and severity of PMIH-related demand. The amount of time spent dealing with PMIH-related incidents is both substantial and disproportionate relative to other forms of police demand.

Background

Engaging with persons with mental ill-health (PMIH) has become an integral part of, and a significant challenge to, contemporary policing. This has been associated, at least in part, with both a growth and shift in the nature of the demands being placed on police forces. In addition to crime and anti-social behaviour, and increasingly, calls-for-service to the police relate to concerns about public safety and welfare (College of Policing, 2015; Boulton *et al.*, 2017; Charman, 2018; Wuschke *et al.*, 2018). Argued to be a consequence of the austerity agenda, which has resulted in a substantive reduction in the non-policing public services provided by the state (Crawford *et al.*, 2018), the police have become positioned a 'service of first resort, rather than last' for dealing with vulnerable people (Winsor, quoted in HMICFRS, 2018). To date, however, there has been a limited

*Department of Sociology, Manchester Metropolitan University, Manchester, UK
**Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, Netherlands
***School of Engineering, University Of Warwick, Warwick, UK

endeavour to account for the extent to which incidents involving PMIH inform the scale of these policing demands and the resources committed in their address. Addressing this shortfall serves as the motivation for this article.

Given that approximately one-quarter of people in the UK experience a mental health problem each year (Mind, 2020) and that one in six adults have a Common Mental Disorder (McManus *et al.*, 2016), it is to be expected that PMIH will come in to contact with the police. Yet, a disproportionately large number of arrestees and victims suffer from mental ill-health (Teplin *et al.*, 2005; White *et al.*, 2006). Such contact can be habitual and generate sub-optimal outcomes, with PMIH trapped in a continuous cycle of police custody, prison, emergency hospital care, and life on the streets (Akins *et al.*, 2016). Engagement with PMIH can also occur irrespective of whether a crime has taken place, with the police seen as a safety net in the absence of appropriate community support networks, acting to refer PMIH to public health services. This role has been recognized for some time. A study in the early 1970s, for example, found that 40% of the patients voluntarily admitted to hospital for mental ill-health issues in the USA were brought to the institution by the police (Gilboy and Schmidt, 1971).

In recent years, however, the scale and nature of policing engagement with PMIH are perceived to have grown and widened substantially, understood as an unavoidable consequence of the design and funding of social policies and of policing models (Livingston, 2016). In these terms, Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) in England and Wales has expressed concern over the extent to which PMIH-related demand is 'placing an intolerable burden on police officers and staff' (HMIC, 2018, p. 3). In short, police forces perceive that they lack the resources (both capacity and capability) required to address the demand associated with PMIH. Comparable concerns to those in England and Wales have been raised in the USA (National Council of State Governments, 2002) and the Netherlands (O'Hagan and Quinn, 2018).

In this light, calculating the scale of calls-for-service entailing engagement with PMIH has become a priority in evidence-based policing research and practice. Using police-recorded data, it has been estimated that around 2% of emergency calls for service relate to emotionally disturbed persons (EDPs; Vaughan *et al.*, 2016) or PMIH (Hartford *et al.*, 2005) in Canada. In the USA, this has been calculated as being as high as 7% in large police jurisdictions, with 92% of police officers claiming experience of one such contact in the previous month (Cordner, 2006). A review of 13 different studies, predominantly in North America, found that approximately 1% of police calls for service involved individuals with mental disorders (Livingston, 2016). In England and Wales, around 3% of all police-recorded incidents (including non-criminal incidents) have been identified as entailing engagement with PMIH (HMIC, 2018). It is uncertain, however, as to the degree of confidence that can be placed in such estimates, particularly those that rely on officer discretion (Hartford *et al.*, 2005). There is a concern that calls-for-service flagged as involving PMIH underestimate the true scale of such incidents. This has been attributed to a lack of training amongst police personnel to enable recognition that an incident entails engagement with PMIH (let alone the difficulty of doing so at such a time) and inconsistent recording practices (i.e. failure to flag an incident with a mental health marker). Moreover, it remains unclear as to what extent such incidents translate into deployed policing resources (e.g. time). Here, evidence to date suggests that incidents entailing engagement with PMIH take longer to resolve than those which do not (Charette *et al.*, 2014). In England, decision models have been used to estimate the average annual cost to public health and police per PMIH incident, but analysis has been limited to a select sample of individuals (Heslin *et al.*, 2017).

Within this context, the research reported here makes two significant contributions enabled through the assessment of the Big Data routinely captured by police forces via a case study of Greater Manchester (UK). First, we deploy a bespoke text mining (artificial intelligence) algorithm to assess the narrative embedded in police incident logs in order to estimate the proportion and nature (severity) of police calls-for-service entailing engagement with PMIH. Second, we use Global Positioning System (GPS) data, which provide a digital spatial and temporal footprint of on-duty police officers, to assess the amount of time the police expend in responding to calls-for-service entailing engagement with PMIH.

Measuring PMIH-related policing demand

The concept of 'police demand' is ill-defined (Laufs et al., 2020). Efforts to qualify it have tended to categorize it as comprising public, protective, and internal demand (see NPCC Performance Management Coordination Committee, 2017; Laufs et al., 2020). Public demand, also referred to as reactive demand, comprises the response to crime and non-crime incidents that have been reported to the police. It is this measure of policing demand that we are interested in here. The existing evidence base exploring PMIH-related policing demands has tended to measure public demand with reference to calls-for-service data. For instance, Vaughan et al. (2016) used calls-for-service over a 1-year period in Vancouver (Canada) and filtered these according to those 'flagged' as involving EDPs. The EDP marker is added by officers at the end of a call based on an assessment of the individual during the incident (White and Ready, 2007). This enables the reporting of a basic descriptive statistic, that is, the proportion of total calls-for-service flagged with an EDP marker. In this case, just over 2% of the total calls-for-service were flagged as involving EDP.

A comparable approach to generating a measure of PMIH-related public demand has been adopted in numerous other studies in Canada (Hartford et al., 2005; Teller et al., 2006; Crocker et al., 2009; Hoch et al. 2009; Vaughan et al., 2017), London, Ontario (Crocker et al., 2009; Hoch et al., 2009), the USA (Kessell et al., 2009), and in England and Wales (HMIC, 2018). As an alternative to this method, others have surveyed proactive officers and asked about the frequency of their experience on PMIH-related incidents, which was then extrapolated to compute an estimate of citywide public demand (see inter alia Panzarella and Alicea, 1997). Providing an overview of this research, Livingston (2016) found that the mean percentage of total calls-for-service attributable to PMIH to be around 1%, though figures can vary considerably with smaller studies reporting up to 20% of calls-for-service as involving PMIH (Holley and Arboleda-Florez, 1988). Here, we note that cross-country comparisons require being caveated by the potential variation in how countries (and jurisdictions) define and flag incidents related to mental ill-health.

A key advantage of using police-recorded incident (calls-for-service) data is that records tend to be geocoded ('where') and time-stamped ('when'). This permits exploration of spatial and temporal variation in PMIH-related police demand and has helped identify 'hotpots' of PMIH incidents, localities, and time periods that account for a disproportionately large volume of calls-for-service (Vaughan et al., 2016; Vaughan et al., 2017). The value of such research is that it can be used to inform the provision of policing and partnership working, for example, ensuring that sufficient officers specializing in mental ill-health, as well as other public health resources, are available at the right time and in the right place.

Nevertheless, the usage of these data to quantify public demand has its shortcomings. First, as identified earlier, police officers and emergency call handlers typically lack the training required to make an accurate and consistent assessment of

whether an incident involves PMIH. Nor is it necessarily possible to do so within the confines of an incident, given the complexity of mental health needs. Secondly, many incidents that involve PMIH are not 'flagged' as such, given that the classification of incidents tends to focus on those characteristics for which the police hold a statutory responsibility to report, for example, crime. Cumulatively, these factors are likely to result in an underestimate of cases involving policing engagement with PMIH (College of Policing, 2015; Livingston, 2016). Thirdly, while calls-for-service data provide an estimate of the scale of PMIH-related demand, they do not shed insight upon the policing resources committed in its address. In this regard, a limited evidence base is suggestive of officers spending longer to manage PMIH-related calls-for-service than they do with other calls-for-service such as traffic incidents, burglaries, and assault (Cordner, 2006). Moreover, engagement with PMIH irrespective of the nature of the call-for-service is perceived to take twice as long and involve more officers than incidents that do not entail engagement with PMIH (Charette *et al.*, 2014). Significantly, officers with expertise in mental health from the Metropolitan Police Service have estimated that around 20% of police time is spent dealing with issues relating to PMIH (Adebowale, 2013). This being said, data-driven estimations of deployed frontline resource, as opposed to estimations of a proportional breakdown of calls-for-service, are lacking. The significance of this shortfall has been highlighted by recent evidence from the USA, which suggests that public health-related incidents occupy a disproportionately large amount of officer time (Ratcliffe, 2021).

Given the current deficiencies in understanding the scale of public demand originating from incidents involving PMIH, and of the subsequent frontline resource committed to meeting this demand, this study holds two key aims. First, to establish a more accurate method for assessing the volume of PMIH-related calls for service, we do this through linking calls-for-service data with incident logs, and by using a bespoke text mining algorithm to automatically, objectively, and consistently identify calls-for-service involving PMIH. We suspect that by harvesting additional information using an objective, computational assessment of text logs that we will unmask a significant amount of PMIH-related demand that has been overlooked by manual flagging. Secondly, to estimate the proportion of frontline or deployed policing resources spent dealing with PMIH-related calls-for-service. We achieve this through linking calls-for-service data with GPS data, the digital spatial, and temporal footprint of officers deployed to calls-for-service.

Data and methods

Data

The data for this study were made available through an information-sharing agreement with Greater Manchester Police (GMP), and following the approval of the University's research ethics committee. The data were transferred, stored, and analysed in alignment with General Data Protection Regulations in a secure computer environment. Two principal data sources are used to meet the aims of this study, namely, police-recorded incident data inclusive of incident logs and GPS data.

Incident data. The incident data used in the analysis are calls-for-service recorded by GMP between January 2014 and September 2017. GMP serves Greater Manchester, a metropolitan county in the North West of England, UK, with a population of around 2.8 million people. The county contains 10 local authorities, each with its own highly urbanized centres, namely, Manchester, Salford, Trafford, Wigan, Bolton, Bury, Rochdale, Oldham, Tameside, and Stockport.

During this time period, a total of 3.65 million incidents were recorded by GMP following emergency calls-for-service. Each incident record contains information about when the call was made

(date and time, to the nearest minute), where it was made from (using Easting and Northing coordinates), a response grade determining the seriousness of the incident, a 'closing code', text log, and if the incident resulted in a crime being committed and recorded, a unique crime number. Of particular relevance to this study are the 'closing codes', the text logs associated with each incident, and the response grades.

Closing codes. Closing codes provide a brief description of the incident and an overview of how it was concluded (HMICFRS, 2017a,b). Descriptions are based on the National Incident Category List, part of the National Standard for Incident Recording (NSIR), which is used by all forces across England and Wales, including GMP. A specific qualifier ('G17') exists in GMP data for incidents deemed to be related to mental health. While the qualifying code (i.e. G17) may differ between forces, the description of the incident as mental health-related is in accord with the NSIR. The qualifier is added based on a final assessment of the information available to both the call handler and officers at the scene. National Police Chief's Council define a mental health incident as 'any police incident thought to relate to someone's mental health where their vulnerability is at the centre of the incident or where the police have had to do something additionally or differently because of it' NPCC (2018). It is through the G17 closing code that it is possible to calculate the proportion of calls-for-service attributable to PMIH as a measure of public demand.

Text logs. Incident text logs contain unstructured narratives for each call-for-service as detailed by the call handler. Logs contain summaries and/or transcriptions of the conversation between the call handler and the caller. They are further supplemented, as the incident unfolds, with information passed on by other members of the public, the officers responding to the incident, and partner agencies (e.g. paramedics). In some instances, text logs will contain historical information about individual(s) involved, which might be relevant to the ongoing incident, such as details passed on by partner agencies (e.g. social services). As noted, there is concern that manually-flagged mental health incident labels miss a considerable number of PMIH-related incidents. Thus, relying on the G17 closing codes in isolation is likely to underestimate the scale of such demand. We suspect that the incident text logs will contain information that, when appropriately harvested, will assist in identifying these missing cases. In this study, the incident logs are used to conduct the text mining analysis, with the aim of uncovering PMIH-related demand which may not have been flagged with the relevant closing code (G17).

Response grades. Response grades are determined by the call handler and based on the threat, risk, and harm posed to those involved in the incident and to the general public. Response grades serve to manage public expectations of the policing response to an incident and aim to ensure that frontline resources (officers) are appropriately and timeously deployed. There are five response grades. Grade 1 (emergency) is allocated to the most serious incidents, which requires police attendance within 15 min of the call being received. Grade 2 (priority) require attendance within 1 h, Grade 3 (routine) require attendance within 4 h, and Grade 4 (scheduled) require attendance or a resolution within 48 h. The least serious, Grade 5 (telephone), does not require police attendance and can be resolved over the telephone or by other means. These data can be used to assess the seriousness of typical PMIH-related incidents (an important consideration when considering the extent of frontline resources consumed by such incidents) and whether the reliability of police-flagged incident records varies according to the seriousness of a PMIH-related incident.

GPS data. The GPS data comprise the digital spatial and temporal footprint of on-duty GMP officers, via their Airwave radios, who are deployed to incidents by emergency call handlers between October 2016 and September 2017. Officers, along with assets (e.g. vehicles), are

linked to an individual incident record from their initial deployment to its resolution, providing a measure of the resources committed in the address of PMIH-related incidents. A time-stamp details when (to the nearest second) an officer (or officers) was assigned to an incident by a call handler and when that officer confirmed that their assistance was no longer required or had been redeployed to another incident.

Methods

In line with the aims of the study, the analysis was conducted in two stages.

Text mining. In order to establish a more accurate measure of the scale of PMIH-related calls-for-service, we deployed a bespoke text mining algorithm on the raw incident text logs. This was motivated by the recognition that the current practice of manually flagging incidents for mental health using the G17 closing code is likely to be inconsistent and generate an underestimate of the volume of PMIH-related incidents. The algorithm, designed to search incident text logs for references to mental health, was developed in three principal stages (for full details, see Haleem *et al.*, 2019). First, a training dataset was generated comprising incidents randomly sampled from May 2014 ($N = 400$). Secondly, using the corresponding police-generated text logs, these incidents were manually reviewed and classified by two researchers specializing in text mining methodologies, including those used specifically for identifying mental health issues, and an active officer in GMP with frontline experience in dealing with PMIH-related incidents. Classifications were conducted independently and verified using Cohen's Kappa coefficient for inter-rater reliability (see Cohen, 1960) which indicated a consistent and reliable flagging of PMIH incidents (between 0.63 and 0.72). Thirdly, a library of key mental health-related words and phrases extracted from the training dataset was used to teach a Convolutional Neural Network (CNN) to automatically identify

PMIH-related incidents from text logs. This included, in some cases, the identification of the G17 closing code in the text log, rather than the incident being formally flagged as such in the database. The CNN was fine-tuned iteratively and verified in comparison to the training dataset. Cross-validation checks indicated a 90% level of agreement between police-flagged incidents and those flagged by the text mining algorithm. We have included examples of words and phrases used to flag incidents in the Appendix 1, but for further methodological details, we refer readers to the corresponding write-up (see Haleem *et al.*, 2019).

The CNN was then deployed on the 3.65 million raw incident text logs spanning between January 2014 and September 2017 to automatically flag incidents associated with PMIH. The proportion of incidents identified through text mining were then compared to the proportion of police-flagged incidents both in sum and by response grade, thus opening prospect of assessing the degree to which police-recorded incident data capture a reliable measure of PMIH-related demand in general, and by the seriousness (response grade) of incidents in particular.

To assess the comparability of the hotspots generated by the police-flagged and the CNN-flagged incident data, we visualized their spatial patterning. To this end, Local Moran's I, a standard Local Indicator of Spatial Association (LISA; Anselin, 1995), was deployed on incident counts aggregated to Lower Super Output Areas (LSOAs). Analysis was conducted in GeoDa version 1.12 using queen continuity. *P*-values were calculated using 999 randomized permutations with a 'significance' threshold of <0.05. Greater Manchester comprised 1,673 LSOA, a statistical geography approximating a neighbourhood, each housing approximately 1,600 people. Findings from the LISA analysis provide an overview of the degree to which incidents cluster in space (e.g. high counts surrounded by high counts, high counts surrounded by low counts). The maps are visualized using hexograms (Harris *et al.*, 2018) which have been shown to

maintain patterns of spatial clustering whilst ensuring that each areal-unit is given equal emphasis in the visualization (Langton and Solymosi, 2019).

Time computation. The incidents identified using the bespoke CNN were then linked to the GPS data using a unique identifier. This restricted the study period for computing frontline policing deployment to October 2016 to September 2017. The time spent by each officer (as identified by their unique collar number) at an incident was calculated based on the difference between the 'start' timestamp, when officers were allocated to that incident (thus including the journey-to-incident) and the 'end' timestamp, when officers were redeployed or had resolved the incident. These data were then aggregated by incident to ensure that the time computation was cumulative. The time resources per incident were computed to the nearest minute and a set of descriptive statistics generated to summarize the central tendency and spread of deployment times for PMIH-related incidents flagged by the CNN. Again, these figures are reported by response grade.

Results

The findings are reported in line with the principal aims of the study.

The scale of PMIH-related policing demand

Of the 3.65 million incidents recorded by GMP during the study period, just over 70,000 were flagged with the G17 mental health qualifier or closing code, comprising 1.9% of the total incidents. This finding is consistent with proportions reported by other police forces in England and Wales, including the Metropolitan Police Service, Yorkshire, Dorset and Avon, and Somerset, as well as the aggregate nationwide figure (College of Policing, 2015). It is also comparable with the findings of studies conducted in North America (Vaughan *et al.*, 2016; Livingston, 2016). In sharp

contrast, however, an additional 290,000 incidents were classified as PMIH-related by the text mining algorithm. In total, therefore, 9.9% of all incidents were PMIH-related. In these terms, it can be concluded that existing recording practices serve to grossly underestimate PMIH-related policing demand.

PMIH-related incident hotspots

In addition to assessing differences in scale of PMIH-related demand, it is also important to examine evidence of their spatial manifestation, as this can serve to guide the targeting of services. We do so here, with a specific focus on exploring whether a distinction exists between the spatial patterning of PMIH-related calls-for-service identified by closing codes and the larger proportion of such incidents identified by the text mining algorithm. Here, we note that around 5% of incidents recorded by GMP occurred outside the boundaries of Greater Manchester or lacked precise geographic information, and have therefore been excluded. Figure 1 serves to visualize the findings of this exercise. The maps present clusters, or hotspots, of LSOAs, which are characterized by similar and/or dissimilar incident counts. The 'High-High' classification denotes LSOAs with a high number of PMIH-related incidents (relative to the mean) which are spatially proximate to LSOAs with similarly high PMIH-related counts. The localities characterized by High–High clustering tend to occupy the urban centres of Greater Manchester, that is, Manchester and Salford in the centre of the map, Wigan to the west, Bolton and Bury to the north west, Rochdale and Oldham to the north east, and Tameside and Stockport in the south east.

In overview, these findings imply a substantial similarity between the spatial clustering, or hotspots of PMIH-related incidents derived by both methods. In other words, despite the distinction in the volume of incidents identified by each method, this does not serve to generate significant variance in the appreciation of the spatial
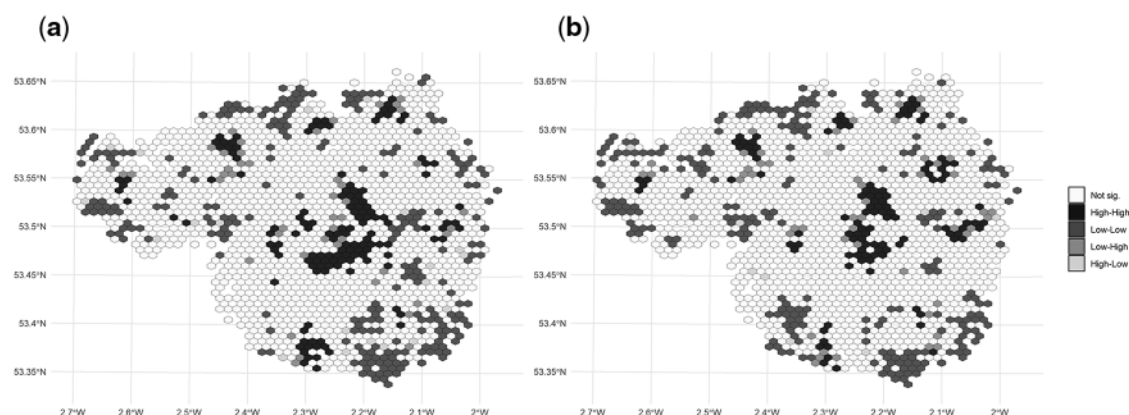
Figure 1: Local Moran's I maps for comparing the spatial patterning of PMIH-flagged incidents using (**a**) closing codes and (**b**) text mining

manifestation of PMIH-related calls-for-service. That said, there are fewer PMIH-related hotspots derived from the text mining algorithm, implying a sharper or more dense spatial patterning, particularly in the centre of the map (Manchester).

Here, we note that the identification of 'high' and 'low' clusters is relative to the number and mean of incident counts in each dataset. For instance, there are High-High clusters present in the closing codes that are not present in the text mining map (e.g. in Manchester city centre), despite there being more incidents overall, simply because these areas are no longer considered 'high' relative to other LSOA in the data. The raw count data for each dataset are reported in the Appendix 2, which confirms that spatial distribution of counts is comparable between datasets despite the disparity in absolute numbers.
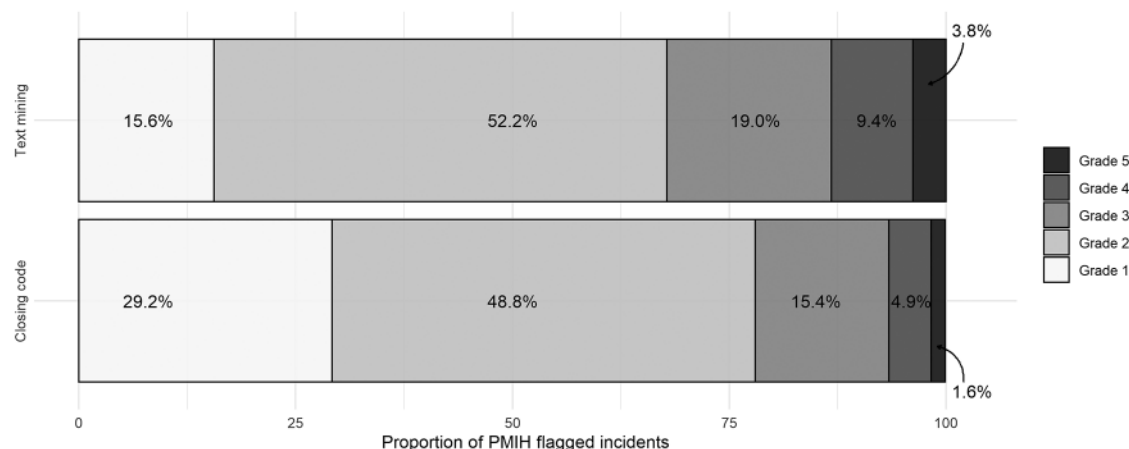
PMIH-related demand by response grade

Table 1 below reports the proportion of total incidents by response grade flagged as PMIH-related by police closing code and by the text mining algorithm. Via both measurement procedures, PMIH-related incidents are most likely to be graded as 1 or 2, implying that they tend being serious in nature. Given that 13.8% of all incidents classified as requiring an emergency response (Grade 1) and 19.9% of all incidents classified as requiring a priority response (Grade 2)

are PMIH-related, as identified by the text mining algorithm, it is clear that a substantial proportion of frontline policing resources are devoted to engaging with PMIH. This significant insight is masked by existing police recording (closing code) practices.

To further clarify these insights, Figure 2 visualizes the distribution of response grades for PMIH-related calls-for-service only and according to whether such incidents were manually flagged or identified via the text mining algorithm. In overview, this confirms that PMIH-related calls-for-service are overwhelmingly judged as being serious in nature and as requiring a rapid policing response. Of those incidents flagged as PMIH-related via closing code, 78.0% were either classed as either an emergency (Grade 1) or a priority (Grade 2). For those identified via the text mining algorithm, this figure was lower but was still high at 67.8%. In contrast, only 6.5% of PMIH incidents flagged via a closing code were given a scheduled or telephone response (Grades 4 and 5), whereas the equivalent figure for those flagged using text mining was 13.2%.

The frontline resource deployed to PMIH-related calls-for-service

Given the scale of engagement with PMIH, identified in the study period January 2014 to

Figure 2: Breakdown of response grades for PMIH-related incidents flagged by closing codes and the text mining algorithm, respectively

September 2017, a key question becomes as to how this translates into the deployment of frontline policing resources. To examine this issue, we link the PMIH-related incident records, identified through the application of the text mining algorithm, to the GPS data of individual officers deployed to these incidents. The availability of the GPS data restricts this exercise to the period between October 2016 and September 2017. During this time, 102,000 incidents were identified as involving engagement with PMIH. Of these, approximately 65,000 (64%) held corresponding GPS data, implying that the remainder (36%) were concluded without frontline policing resource being deployed. The absence of frontline policing deployment can be attributed, in part, to incidents of a less serious nature. Thus, only 6.6% of PMIH-related incidents categorized for telephone resolution (Grade 5) held linked deployment data. In contrast, it was possible to link GPS records to 89.4% of emergency response incidents (Grade 1). Nevertheless, it is significant that the percentage of incidents with corresponding GPS data are slightly lower than that which might be expected given the initial grading of PMIH-related incidents, as reported above. This may represent a weakness in the data or, following the initial grading of an incident, other policing demands served

to inhibit the availability of frontline line resource and/or other agencies served to act resolve the incident.

Table 2 reports a set of descriptive statistics detailing the time (in minutes) that police officers spent responding to PMIH-related incidents, by response grade. These findings are based on the cumulative frontline policing resource deployed to incidents. In other words, they serve to capture situations in which multiple officers were deployed to the same incident. The existence of high outliers, cases in which a disproportionately large amount of frontline policing resource were required to resolve, served to skew the calculation of the mean resource allocated to incidents. In these terms, it is more meaningful to report the median number of minutes attributable to each response grade. Thus, emergency response incidents (Grade 1) tended to consume a total of 151 min (or 2.5 h) of cumulative officer time, priority response incidents (Grade 2) 125 min (or 2 h), routine response incidents (Grade 3) 105 min, and scheduled response incidents (Grade 4) 112 min. The least time-consuming response grade was that of telephone resolution (Grade 5), which tended to consume 103 min of officer time. This ranking of deployed resource by incident grade was consistent across both the lower and upper quartiles of the data. In

Table 1: The proportion of all incidents by response grade flagged as PMIH-related by police closing code and by the text mining algorithm

Response grade	Closing code (%)	Text mining (%)
1: Emergency response attendance within 15 min of call receipt	5.0	13.8
2: Priority response attendance within 1 h of call receipt	3.6	19.9
3: Routine response attendance within 4 h	1.2	7.5
4: Scheduled response attendance or other resolution within 48 h	0.5	5.1
5: Telephone resolution first-time telephone resolution of a call	0.2	2.0
PMIH incidents as a proportion of all incidents	1.9	9.9

Table 2: Descriptive statistics of the cumulative time (in minutes) spent on PMIH incidents by response grade

Response grade	Mean	Response time (in minutes)				
		(Q1) 25%	Median	(Q3) 75%	Min.	Max.
1: Emergency response attendance within 15 min of call receipt	296.5	76.0	151.0	308.0	0	26,452
2: Priority response attendance within 1 h of call receipt	251.6	63.0	125.0	236.0	0	132,568
3: Routine response attendance within 4 h	261.1	46.0	105.0	220.0	0	250,604
4: Scheduled response attendance or other resolution within 48 h	412.9	44.0	112.0	246.5	0	27,177
5: Telephone resolution first-time telephone resolution of a call	598.2	22.0	102.5	244.3	0	20,116
All PMIH incidents	269.3	63.0	128.0	250.0	0	250,604

Note: Zero minimum values arise as a consequence of rounding.

aggregate, PMHI-related incidents consumed around 2 h (128 min) of police officers' time. The most significant finding, however, is that 19.5% of police resources were allocated to incidents requiring engagement with PMIH. In other words, and between October 2016 and September 2017, around one-fifth of the total volume of deployed policing frontline policing resources was spent dealing with incidents involving PMIH. Given that 9.9% of all calls-for-service were

identified as PMIH-related, this finding suggests that PMIH-related incidents consume a substantial and disproportionate proportion of frontline policing resources.

Discussion

That police records serve to underestimate the scale and misrepresent the seriousness of PMIH-related public (reactive) demand aligns

with concerns surrounding the capability of existing recording procedures to capture the complexities shaping the demand for service. Whilst the 1.9% of cases identified via police closing codes are consistent with findings reported elsewhere in England and Wales, noting the caveats in comparing cross-nationally (see *inter alia* Vaughan *et al.*, 2016), the 9.9% of cases identified via the text mining of incident logs demonstrate that police-recorded data serve to significantly downplay the extent if not severity of public demand associated with PMIH. The finding that 9.9% of incidents are PMIH-related is itself a likely underestimate of the real extent of such demand given the challenge of identifying mental health needs, particularly in the absence of specialist training and the lack of prioritization for doing so. Moreover, we recognize that the additional public demand identified here will follow through into internal demand (e.g. back-office costs). Nevertheless, this research provides an empirical basis to critique existing estimates as well as, hopefully, the impetus to improve existing recording practices. It also illustrates the potential of existing policing data resources and of data science methodologies to contribute to this process. At the least, this finding and approach can serve as a basis to motivate comparative investigations across different study regions.

The study further moved to illuminate the spatial patterning of PMIH-related public demand. It found a reasonable degree of overlap between the PMIH-related incident hot spots identified by manually flagged PMIH police records and those identified by the text mining algorithm, enabling some confidence of those reliant on the former to target services. That said, the PMIH-related hotspots generated a sharper patterning (based on a larger count), enabling a more precise targeting of policing and other public agency resources to the areas of highest demand. On the whole, PMIH-related incident hotspots tended to concentrate in the town and city centres of Greater Manchester, places reflective of a set 'attractors' associated with PMIH such as healthcare facilities and licensed premises (Vaughan *et al.*, 2016). Given the precise geographical and temporal qualities of the data available to police forces, future research might consider yet further assessment of the patterning of PMIH-related demand as a means to develop more spatially and temporally refined interventions.

The finding that PMIH-related incidents account for 19.5% of all policing frontline or deployed officer resources is perhaps the most significant insight of the research. While it is consistent with existing practice-informed perception, as reported to the Independent Commission on Mental Health and Policing (Adebowale, 2013), this is the first study to offer robust empirical evidence in support of this claim. As such, it highlights the significant role that the police currently play in safeguarding public well-being and corroborates the view that the police are confronted with an 'intolerable burden' (HMIC, 2018, p. 3) in delivering this service in the absence of sufficient training and resource to engage with PMIH. A key next step, building on insights emergent from Canada (Vaughan and Andresen, 2018), will be to compute the costs associated with the excess demand identified and to compare these to existing estimates (see Heslin *et al.*, 2017).

Within the current context of policing, this study serves to highlight the need for improved recording practices aligned with enhanced policing capability and capacity to meet the demand for safeguarding as well as the requirement for greater inter-agency collaboration. In Canada, where comparable concerns have been raised over data recording standards for PMIH police contact (Vaughan and Andresen, 2018), recommended remedies extend to include deepening information sharing between police forces and mental health professionals, not only both in terms of knowledge and expertise, but also data (Huey *et al.*, 2021). GMP has made steps to address this by introducing a triage system in which mental health practitioners are embedded in the call centre and via the training of both call handlers and active officers to help identify and then record incidents involving PMIH.

Conclusion

This study has made two key contributions. First, using a bespoke text mining algorithm on police incident text logs in Greater Manchester, we have demonstrated that incidents manually flagged by police vastly underestimate the scale of public demand originating from PMIH. Police records indicate that 1.9% of total incidents involve PMIH, but automated text mining identified 9.9%, unmasking a further 290,000 incidents during the study period. Incidents flagged as involving PMIH tend to be classified as serious, demanding emergency or priority response times. Secondly, using GPS data on police officer responses to PMIH-related incidents, we provide a data-driven estimate of frontline police time and resources allocated to meet this demand. Findings suggest that officers spend around 2 h on such incidents, although this varies by response grade. Broadly speaking, the length of time spent on resolving PMIH-related calls for service increases with the seriousness of the incident. Overall, we estimate that 19.5% of frontline police time is spent resolving incidents involving PMIH. The research findings offer a robust evidence base to support recent calls for enhanced training among police officers and call handlers, and encourage deeper cooperation between police and partner agencies, such as health services. In doing so, police-recorded incident data can offer greater reliability and validity as a measure of public demand, and ultimately, the well-being outcomes of PMIH who come into contact with the police can be improved.

Acknowledgements

References

Adebowale, V. (2013). 'Independent Commission on Mental Health and Policing Report'. London: *Independent Commission on Mental Health and Policing*.

Akins, S., Burkhardt, B. C. and Lanfear, C. (2016). 'Law Enforcement Response to "Frequent Fliers" an Examination of High-Frequency Contacts between Police and Justice-Involved Persons with Mental Illness'. *Criminal Justice Policy Review* 27(1): 97–114.

Anselin, L. (1995). 'Local Indicators of Spatial Association—LISA'. *Geographical Analysis* 27(2): 93–115.

Boulton, L., McManus, M., Metcalfe, L., Brian, D., and Dawson, I. (2017). 'Calls for Police Service: Understanding the Demand Profile and the UK Police Response'. *The Police Journal* 90(1): 70–85.

Charette, Y., Crocker, A. G., and Billette, I. (2014). 'Police Encounters Involving Citizens with Mental Illness: Use of Resources and Outcomes'. *Psychiatric Services* 65(4): 511–516.

Charman, S. (2018). From Crime Fighting to Public Protection: The Shaping of Police Officers' Sense of Role. *Perspectives on Policing: Paper, 3*. London: Police Foundation.

Cohen, J. (1960). 'A Coefficient of Agreement for Nominal Scales'. *Educational and Psychological Measurement* 20(1): 37–46.

College of Policing. (2015). *College of Policing Analysis: Estimating Demand on the Police Service*. Retrieved from College of Policing: https://www.college.police.uk/News/College-news/Documents/Demand%20Report%2023_1_15_noBleed.pdf (accessed 1 July 2021).

Cordner, G. W. (2006). *People with Mental Illness*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

Crawford, R., Disney, R., and Simpson, P. (2018). *The Determinants of Local Police Spending* (No. W18/09). IFS Working Papers.

Crocker, A. G., Hartford, K., and Heslop, L. (2009). 'Gender Differences in Police Encounters among Persons with and without Serious Mental Illness'. *Psychiatric Services* 60(1): 86–93.

Gilboy, J. A. and Schmidt, J. R. (1971). 'Voluntary Hospitalization of the Mentally Ill'. *Northwestern University Law Review* 66: 429.

Haleem, M. S., Han, L., Harding, P. J., and Ellison, M. (2019). An Automated Text Mining Approach for Classifying Mental-Ill Health Incidents from Police Incident Logs for Data-Driven Intelligence. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. Piscataway, NJ: IEEE, pp. 2279–2284.

Harris, R., Charlton, M., and Brunsdon, C. (2018). 'Mapping the Changing Residential Geography of White

British Secondary School Children in England Using Visually Balanced Cartograms and Hexograms'. *Journal of Maps* **14**(1): 65–72.

Hartford, K., Heslop, L., Stitt, L., and Hoch, J. S. (2005). 'Design of an Algorithm to Identify Persons with Mental Illness in a Police Administrative Database.' *International Journal of Law and Psychiatry* **28**(1): 1–11.

Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services. (2017a). *State of Policing 2016: The Annual Assessment of Policing in England and Wales.*

Her Majesty's Inspectorate of Constabulary and Fire and Rescue Services. (2017b). *Crime Data Integrity Programme: Crime-recording Process.* https://www.justiceinspectorates.gov.uk/hmicfrs/our-work/article/crime-data-integrity/crime-recording-process/ (accessed 22 October 2020).

———. (2018). Policing and Mental Health: Picking Up The Pieces. *Retrieved from HMICFRS: https://www.justiceinspectorates.gov.uk/hmicfrs/wp-content/uploads/policing-and-mental-health-picking-up-the-pieces.pdf (accessed 1 July 2021).*

Heslin, M., Callaghan, L., Barrett, B. et al. (2017). 'Costs of the Police Service and Mental Healthcare Pathways Experienced by Individuals with Enduring Mental Health Needs'. *The British Journal of Psychiatry* **210** (2): 157–164.

Hoch, J. S., Hartford, K., Heslop, L., and Stitt, L. (2009). 'Mental Illness and Police Interactions in a Mid-Sized Canadian City: What the Data Do and Do Not Say'. *Canadian Journal of Community Mental Health* **28**(1): 49–66.

Holley, H. L. and Arboleda-Florez, J. (1988). 'Criminalization of the Mentally Ill: Part I. Police Perceptions'. *The Canadian Journal of Psychiatry* **33**(2): 81–86.

Huey, L., Ferguson, L., and Vaughan, A. D. (2021). 'The Limits of Our Knowledge: Tracking the Size and Scope of Police Involvement with Persons with Mental Illness'. *FACETS* **6**(1): 424–448.

Kessell, E. R., Alvidrez, J., McConnell, W. A., and Shumway, M. (2009). 'Effect of Racial and Ethnic Composition of Neighborhoods in San Francisco on Rates of Mental Health-Related 911 Calls'. *Psychiatric Services* **60**(10): 1376–1378.

Langton, S. H. and Solymosi, R. (2019). 'Cartograms, Hexograms and Regular Grids: Minimising Misrepresentation in Spatial Data Visualisations'. *Environment and Planning B: Urban Analytics and City Science, 48*(2): 348-357.

Laufs, J., Bowers, K., Birks, D., and Johnson, S. D. (2020). 'Understanding the Concept of 'Demand'in Policing: A Scoping Review and Resulting Implications for Demand Management'. *Policing and Society,* 1–24. DOI: 10.1080/10439463.2020.1791862.

Livingston, J. D. (2016). 'Contact between Police and People with Mental Disorders: A Review of Rates'. *Psychiatric Services* **67**(8): 850–857.

McManus, S., Bebbington, P., Jenkins, R., and Brugha, T. (2016) *Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014.* Leeds, England: NHS Digital.

Mind. (2020). *Mental Health Facts and Statistics.* https://www.mind.org.uk/information-support/types-of-mental-health-problems/statistics-and-facts-about-mental-health/how-common-are-mental-health-problems/ (accessed 15 December 2020).

National Council of State Governments. (2002). *Criminal Justice and Mental Health Consensus Project.* New York: Council of State Governments.

NPCC Performance Management Coordination Committee. (2017). *Better Understanding Demand – Policing the Future.* Retrieved from NPCC: *https://www.npcc.police.uk/2017%20FOI/CO/078%2017%20CCC%20April%202017%2024%20Better%20Understanding%20Demand%20Policing%20the%20Future.pdf (accessed 1 July 2021).*

NPCC (2018). National Strategy on Policing and Mental Health. Retrieved from NP CC: https://www.npcc.police.uk/Mental%20Health/Nat%20Strat%20Final%20v2%2026%20Feb%202020.pdf (accessed 1 July 2021).

O'Hagan, A. and Quinn, L. (2018). 'Mental Health a Policing Challenge for the United Kingdom and The Netherlands'. *Forensic Research & Criminology International Journal* **6**(2): 00184.

Panzarella, R. and Alicea, J. O. (1997). 'Police Tactics in Incidents with Mentally Disturbed Persons'. *Policing: An International Journal of Police Strategies & Management* **20**(2): 326–338.

Ratcliffe, J. H. (2021). 'Policing and Public Health Calls for Service in Philadelphia'. *Crime Science* **10**(1): 1–6.

Teller, J. L., Munetz, M. R., Gil, K. M., and Ritter, C. (2006). 'Crisis Intervention Team Training for Police Officers Responding to Mental Disturbance Calls'. *Psychiatric Services* **57**(2): 232–237.

Teplin, L. A., McClelland, G. M., Abram, K. M., and Weiner, D. A. (2005). 'Crime Victimization in Adults with Severe Mental Illness: Comparison with the National Crime Victimization Survey'. *Archives of General Psychiatry* **62**(8): 911–921.

Vaughan, A. D. and Andresen, M. A. (2018). 'The Cost of Mental Health-Related Calls on Police Service: Evidence from British Columbia'. In R. J. Mitchell and L. Huey (Ed.), *Evidence-Based Policing: An Introduction.* (pp. 173–185). Great Britain, Policy Press.

Vaughan, A. D., Hewitt, A. N., Andresen, M. A., and Brantingham, P. L. (2016). 'Exploring the Role of the Environmental Context in the Spatial Distribution of

Calls-for-Service Associated with Emotionally Disturbed Persons'. *Policing: A Journal of Policy and Practice* **10**(2): 121–133.

Vaughan, A., Wuschke, K. E., Hewitt, A. N. et al. (2017). 'Variations in Mental Health Act Calls to Police: An Analysis of Hourly and Intra-Week Patterns'. *Policing: An International Journal* **41**: 58–69.

White, M. C., Chafetz, L., Collins-Bride, G., and Nickens, J. (2006). 'History of Arrest, Incarceration and Victimization in Community-Based Severely Mentally Ill'. *Journal of Community Health* **31**(2): 123–135.

White, M. D. and Ready, J. (2007). 'The TASER as a Less Lethal Force Alternative: Findings on Use and Effectiveness in a Large Metropolitan Police Agency'. *Police Quarterly* **10**(2): 170–191.

Wuschke, K. E., Andresen, M. A., Brantingham, P. J., Rattenbury, C., and Richards, A. (2018). 'What Do Police Do and Where Do They Do It?' *International Journal of Police Science & Management* **20**(1): 19–27.

Appendix 1

Examples of keywords and phrases with high classification power in the text mining algorithm. For full technical details, we refer readers to Haleem *et al.* (2019). g17, g16g17, g16, mental, powers, health, mental health, no powers, powers used, no powers used, present, suicidal, hospital, issues, suffers, mn, ambulance, distress, gp, taken, depression, code g16, medication, health issues, code g16g17, mental health issues, amb, dr, d51, medium, mn issues, centre, getting, support, and distress disturbance.

Appendix 2
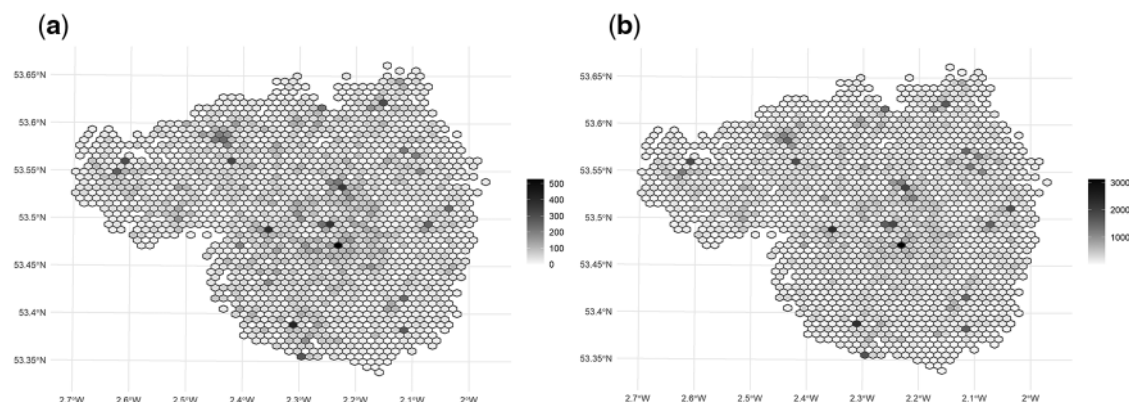
Figure A1: Raw count maps for comparing the spatial patterning of PMIH-flagged incidents using (a) closing codes, and (b) text mining.