SIA
SOCIAL INVESTMENT AGENCY
ORANGA TANGATA

Investing in what works for better lives

Using integrated data to understand mental health and addiction conditions

August 2019

Technical Guide

1908.DP.TGU.MHA

New Zealand Government

Creative Commons Licence

Liability

Intended audience

This guide is written for analysts and data scientists to support data analytics.

Citation

Social Investment Agency 2019. Using integrated data to understand mental health and addiction conditions. Technical Guide. Wellington, New Zealand.

ISBN 978-0-473-49236-6

Published in 2019 by
The Social Investment Agency
Wellington, New Zealand

Acknowledgements

The Social Investment Agency (SIA) would like to acknowledge the support and assistance of all those who participated in the development of this report:

- the Analytical Services team and the National Collections and Reporting teams at the Ministry of Health for their expert advice on the quality and use of the health data
- the Ministry of Justice for their collaboration on data analysis
- Statistics NZ for the use and support of the Integrated Data Infrastructure (IDI). Without the use of this rich data taonga our work would not be possible
- Dr Tom Love, Director, Sapere Research Group Limited for his expert advice and guidance of the data, interpretation of results, and application to policy settings
- Dr Anthony Duncan, National Advisor, Intellectual Disability (Compulsory Care and Rehabilitation) for sharing his knowledge about working with patients and knowledge of pharmaceuticals.

Reviewers

SIA would like to acknowledge the reviewers who provided feedback on early drafts in 2017 or 2018, noting that there has been substantial revision of the content over time:

- Dr Doone Winnard, Counties Manukau DHB
- Ann Sears, Counties Manukau DHB
- Wing Cheuk Chan, Counties Manukau DHB
- Dr Helen Lockett, The Wise Group
- Dr Michael Ryan, Sapere Research Group
- Steven Johnston, Ministry of Health
- Catherine Coates, Ministry of Health
- Catherine Inder, Ministry of Health
- Alice McCallum, Ministry of Health
- Bex Sullivan, Ministry of Health

Contributors to this report

- Kylie Reiri
- Conrad McCormick
- Vinay Benny
- Ernestynne Walsh
- Craig Wright
- Doug Lambert
- Shrividya Ravi
- Akilesh Chokkanathapuram Vasudevan
- Simon Anastasiadis

- Atawhai Tibble
- Uekaha Douglas
- Fiona Thomson
- Marianna Pekar
- Kate Kolich

Integrated Data Infrastructure (IDI) Disclaimer

The results in this report are not official statistics; they have been created for research purposes from the IDI managed by Statistics NZ. The opinions, findings, recommendations, and conclusions expressed in this report are those of the author(s), not Statistics NZ, or other government agencies.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. SIA researchers only had access to anonymised data. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this report have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security, and confidentiality issues associated with using government administrative and survey data in the IDI. Further detail can be found in the privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes.

Any person who has had access to the unit record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

Contents

A new resource for mental health and addiction research in the IDI is available

The Social Investment Agency (SIA) has developed a tool that consolidates Integrated Data Infrastructure (IDI) data about mental health and addiction (MHA) conditions. This tool enables analysis to be undertaken using consistent metrics drawn from data currently available in the IDI.

Identifying the unmet needs of individuals living with mental health and addiction conditions is a key priority of the Mental Health Inquiry (2018) and the current Government.

"Our purpose was to identify unmet needs and develop recommendations for a better mental health and addiction system for Aotearoa New Zealand" NZ Mental Health Inquiry 2018[1]

Service use observed in administrative microdata is one starting point for understanding MHA conditions. This data exists within specific agencies and has previously been accessible only to the staff of those agencies. However, the development of the IDI by Statistics New Zealand has provided a place where data from a range of agencies has been consolidated and linked at an individual level.

The existence of the IDI presents a significant opportunity for detailed research into mental health and addiction conditions. A key barrier to the realisation of such research is the varied presentation of the different data sources within the IDI. This results in a high level of initial effort to identify and combine data prior to undertaking analysis. To address this, SIA together with domain experts, have consolidated five key sources of mental health and addictions data in the IDI into a single table of events.

The table of events is intended as a resource for mental health and addiction researchers. It is being made publicly available for use by analysts to easily define cohorts or populations that have had interactions with MHA-related services. It will provide a starting point for establishing standards for mental health and addictions research and definitions within the research community.

We provide two products intended to be used together:

1. Code to generate the table of events and accompanying researcher instructions - available at SIA's GitHub page[2]
2. Technical documentation - provided by this document, which covers:
1. The health system and points of data collection
2. A description of the consolidated table
3. The quality and reliability of the data
4. A discussion of the suitability of the data for different purposes

[1] https://mentalhealth.inquiry.govt.nz/

[2] https://github.com/nz-social-investment-agency/mha_data_definition

5. Supporting appendices.

Statistics NZ "Five Safes" Framework[3]

Statistics New Zealand's "five safes" framework also applies to mental health data being accessed through IDI. The "five safes" conditions are as follows:

1. Safe people – researchers are vetted and must commit to use data safely
2. Safe projects – projects must be demonstrated to have be in public interest
3. Safe settings – range of privacy and security arrangements are put in place to keep data safe
4. Safe data – data that is available to researchers is de-identified, access to data is limited to what is necessary for the research
5. Safe output – all information is checked to ensure it does not contain any identifying results.

Data Protection and Use Policy[4]

SIA is committed to the safe, ethical and transparent use of social sector data. We are working with the social sector to collectively develop a policy for anyone working with personal data and information (including information that has been de-identified or anonymised). The Data Protection and Use Policy (DPUP) is a collection of principles and guidance that will be supported by tools to enable everyone to easily understand what is appropriate, what is not, and how to work safely and respectfully with people's personal information. DPUP will help to build trust and confidence in the social sector's use of people's data and information, and the quality of that data so that better outcomes can be achieved from it. It is important to know that this policy promotes best practice but will not change legislation.

[3] https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure#five-safes

[4] https://www.sia.govt.nz/investing-for-social-wellbeing/data-protection-and-use/

What questions can available data answer?

Making the best possible use of data and analytics, as well as other evidence, helps the social sector to get the right help to the people who need it. For MHA conditions, particular concerns revolve around understanding unmet need. A sample of the questions that can be answered through this schema is listed in Table 1.

Table 1 Sample questions that can be answered with MHA events (and other administrative data where needed)

High level question	The IDI can help explore these questions
Who needs support?	What macro-level groups appear to need mental health services, but are not receiving them? • *Are Māori living in Gisborne under-serviced (mental health)? (Comparing service uptake against known estimates of prevalence)*
When is the best time to receive support?	What are common life courses for people who use services or treatment (in the IDI)? • *Do people who are referred to specialist services from appropriate services (primary services or school) have better outcomes compared to people who are referred from crisis services (Justice or ED)? (Comparing non-MHA outcomes e.g. employment, recidivism)*
Who accesses services or treatments?	Who received mental health or addiction services or treatments (available in the IDI) and what are their characteristics? • *How old were people when they received services (in the IDI) for the first time?* • *Are services (in the IDI) delivered equitably across socio-economic status?* What experiences do people have before or after an event or before or after receiving a service or treatment? • *What government interactions do people have before they self-harm?* • *What mental health services or treatments do people receive before and after offending?*
How well do services or treatments work for the people accessing them (impact evaluation)?	What is the impact of services and treatments (in the IDI) on non-mental wellness outcomes? • *Do services delivered by NGOs result in more jobs for the people receiving them, compared with services delivered by DHBs?*
How do mental health service access, uptake and outcomes differ for Māori, Pacific peoples, migrants and refugees?	Does the impact of services and treatments (in the IDI) differ depending on ethnic and migrant background? • *How does mental health related service receipt differ by ethnicity?* • *Are services delivered by DBHs result in better or worse health outcomes for Māori than for the rest of the population receiving them?*

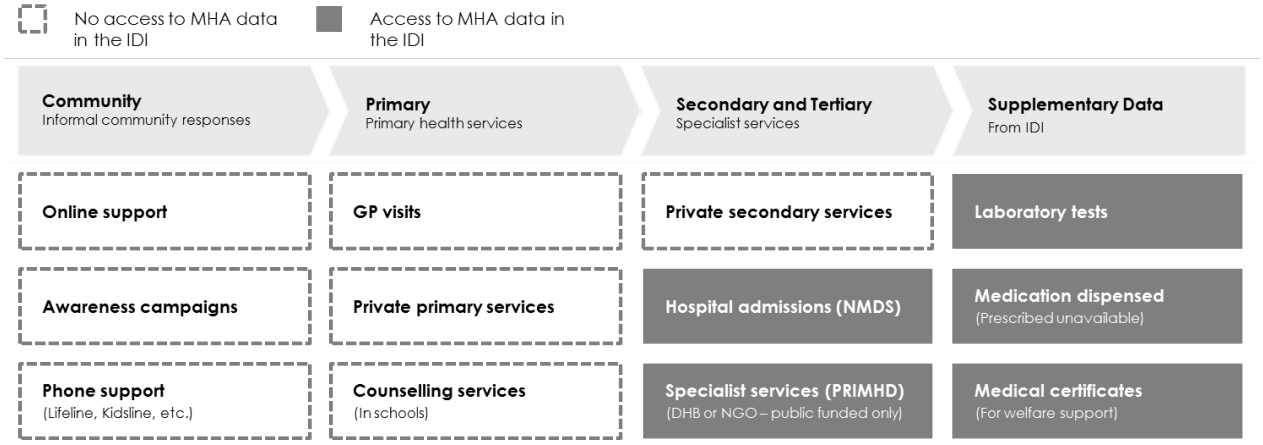
Mental health and addiction events are only observed at specific points of service interaction

An understanding of the origins of the data is required to interpret and apply it correctly. We discuss the different points in the system from which data that captures MHA conditions has been collected and included in the IDI.

MHA data is limited to specialist and supporting services

The health service in NZ is divided into three tiers: community, primary and secondary/tertiary services. Community services capture a significant amount of informal help—accessible to people without the formality of a doctor's appointment. The middle tier—primary services—captures the typical entry point into the public healthcare system. While patients can self-refer straight to secondary (specialist) services, it is also common for a referral to come from a primary service.

Figure 1: Administrative data available in the IDI of individuals who have accessed mental health and addiction services or treatments



Only two parts of the health service shown above are captured in the IDI and provide clearly coded mental health and addiction events. These are specialist services as recorded by the Project for the Integration of Mental Health Data (PRIMHD), and hospital admissions as recorded by National Minimum Dataset (NMDS).

In addition to the tiers shown above, there are three supporting services that cut across different service tiers and types, and provide information about mental health and addiction conditions. These are laboratory tests, dispensed medications, and medical certificates used for welfare support. People interacting with any tier of the health system can interact with any of the supporting services.

Table 1 shows the five datasets identified above, and incorporated into the consolidated table by SIA. Note that a person can appear in multiple data sources. For example, a person may appear in

PRIMHD because they consulted with a mental health specialist and appear in pharmaceuticals dispensing as they have collected a prescription. These data sets need to be considered together.

Table 2: Datasets used to identify population groups who access mental health and addiction services

Service or treatment	Dataset	Dataset description
Hospital admissions	NMDS (publicly funded discharges only)	The National Minimum Dataset is a national collection of publicly and privately funded hospital discharge information, including clinical information, for inpatients and day patients. Privately funded hospital events have been excluded from the version of this dataset in the IDI because it is incomplete.
Specialist services	PRIMHD	The Project for the Integration of Mental Health Data is a Ministry of Health single national mental health and addiction information collection of service activity and outcomes data for health consumers. The data is collected from DHBs and NGOs organisations.
Laboratory tests	Laboratory Claims Collection	The Laboratory Claims Collection contains claim and payment information for laboratory tests.
Medication dispensed	The Pharmaceutical Collection (dispensing data)	The Pharmaceutical Collection supports the management of pharmaceutical subsidies. It is jointly owned by the Ministry of Health and PHARMAC. The Pharmaceutical Collection contains claim and payment information from pharmacists for subsidised dispensed medication through community pharmacies. It includes cancer therapies obtained from hospital pharmacies but no other medications provided through hospital pharmacies.
Medical certificates for welfare support	MSD incapacity data	The MSD incapacity dataset identifies people who receive Job Seeker Health Conditions and Disability (JS-HCD) or Supported Living Payment (SLP). This dataset records the reason for their incapacity to work through incapacity codes, which allows this analysis to identify patients who have received a medical certificate application or renewal as a proxy for primary care.

Data is focused on service use

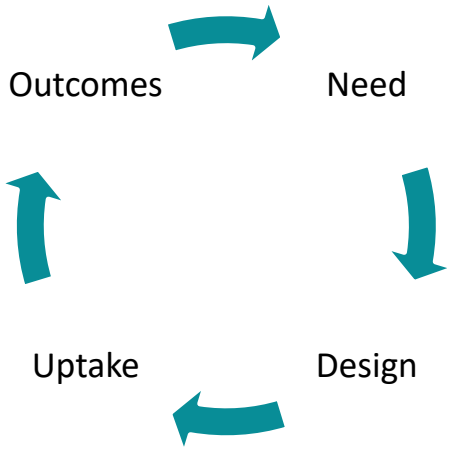
Effective service delivery requires the identification of need, design of a service model to address the need, delivery and uptake of the service, and positive outcomes to arise from service interaction. It follows that research into mental health and addiction conditions will want to understand these different stages.

There is no guarantee that individuals receiving treatment experience positive outcomes. Differences in uptake and outcomes highlight *unmet needs* in:

- those with need who do not access treatment
- those who access treatment without any lasting outcomes.

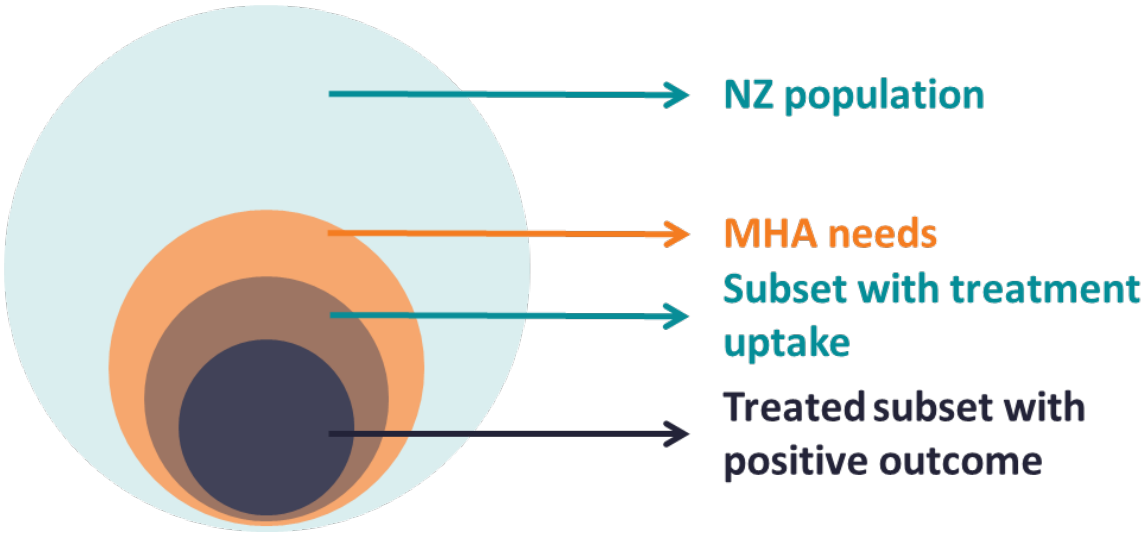
Figures 2 and 3 illustrate these dynamics.

Figure 2: Stages of effective service delivery



Healthcare providers aim to address identifed need with by designing and providing treatments. The provision of a service/treatment aims to improve wellbeing for the treated individuals. The assumption that individuals in need will access the system, and actively participate in treatment – does not hold for many people.

Figure 3: Mental health and addiction need in the population (Not to scale)



Identification of the prevalence of mental health and addiction need has been primarily done via surveys like the Ministry of Health's Te Rau Hinengaro: The New Zealand Mental Health Survey (Browne et al, 2006). While Te Rau Hinengaro is not available in the IDI, a range of Statistics New Zealand surveys are available. These include the Survey of Family, Income and Employment (SoFIE)[5] and the New Zealand General Social Survey (NZGSS)[6] which explicitly ask questions relating to mental health and wellbeing. Te Kupenga[7] is a more recent addition to IDI, it is Statistics New Zealand's first survey of Māori well-being, recognising practices and well-being outcomes that are specific to Māori culture, for example Te Reo, social connectedness, whānau well-being. It has a module related to physical and mental health.

The data consolidated in the resource SIA has prepared is focused on uptake: the use of mental health and addiction services. This is the data that is most extensive in the IDI and the stage of service delivery that is best covered by administrative data.

Designing a functional schema for the consolidated table

SIA has consolidated mental health and addiction related events into a single table to enable IDI researchers to use these definitions as a foundation for their research. IDI researchers with access to the five input data sources can construct the consolidated table by running the code available on SIA's GitHub page[8].

[5] http://archive.stats.govt.nz/browse_for_stats/income-and-work/Income/sofie.aspx

[6] http://archive.stats.govt.nz/browse_for_stats/people_and_communities/Well-being/nzgss-info-releases.aspx

[7] http://archive.stats.govt.nz/browse_for_stats/people_and_communities/maori/te-kupenga.aspx

[8] https://github.com/nz-social-investment-agency/mha_data_definition

Identifying mental health and addiction conditions in IDI data is complex. Our coding architecture reflects the best recommendations we have been able to combine from the Ministry of Health, District Health Boards (DHBs), non-government organisations (NGOs), other clinicians, and data experts.

Researchers with specialised interests will want to refine the data definitions to better reflect populations or conditions they have a specific interest in. SIA provides this resource under a GNU GPLv3[9] license to enable users to refine and extend it, and encourages users to do so. We encourage users to upload their variants of the consolidated table to GitHub as part of the same repository. This will enable the version control built into GitHub to track the different versions of the code. This document describes Version 1.

The structure of the consolidated table is given in Table 2. We have followed the same format as an existing SIA product: the Social Investment Analytical Layer. This enables further integration with our other tools and resources. Each row in the consolidated table corresponds to a single event for a single person. Data provenance (or the origin of the data) is described by the department, datamart and subject_area columns while service / treatment detail is given by the event_type columns.

Table 3: Data Schema for the core table of MHA events identified in the IDI

Column	Explanation
snz_uid	Unique, anonymised identifier for an individual in the IDI. Links individuals across different datasets
department	Government agency supplying the administrative data e.g. Ministry of Health
datamart	One of the five datasets described in Table 1
subject_area	Default of 'MHA'. Column exists for joining with other types of data to differentiate between domains e.g. Mental Health / Addictions (MHA), Benefits (BEN), Education (EDU)
start_date	Start of service / treatment
end_date	End of service / treatment
event_type	Broad diagnosis class describing mental health or addiction condition e.g. Mood, Anxiety. See the discussion of diagnosis class below
event_type_2	Detail of treatment e.g. specific pharmaceutical dispensed
event_type_3	Not applicable unless PRIMHD data where it describes the 'Team Type'.

Classification rules for diagnosis classes

The Table 3: Data Schema for the core table of MHA events identified in the IDI also includes a broad diagnosis class (in the event_type column). These classes are neither medical diagnoses nor mutually exclusive, but seek to provide a common parlance description of a particular class of

[9] https://www.gnu.org/licenses/gpl-3.0.en.html

mental health and addiction condition. They are intended to assist in organising the event table and to accelerate researchers identifying conditions of interest.

Classification rules that generate the diagnosis classes were developed after consultation with Dr Anthony Duncan. The table below describes the current set of classification rules. Note that the diagnosis categories provided by this definition are not exhaustive, and should be treated as simple high-level inferences made from the IDI data – they are not definitive medical diagnosis information.

Table 4: Derivation of MHA service or treatment access from various data sources available in the IDI

Diagnosis class	PRIMHD[10]	Pharmaceuticals	NMDS – publicly funded hospital discharges	MSD incapacity
ADHD		Chemical ID = 3887, 1809, 3880	ICD-10-AM diagnosis code = F900 ICD-9-CMA-II diagnosis code = 31400, 31401	
Anxiety		Chemical ID = 1166, 6006, 1780	ICD-10-AM diagnosis code = F40-F48 ICD-9-CMA-II diagnosis code = 30000-30015, 3002, 3003, 3005-3009, 3060-3064, 30650, 30652, 30653, 30659, 3066-3069, 30780, 30789, 3080-3091, 30922-30982, 30989, 3099	
Autism			ICD-10-AM diagnosis code = F84 ICD-9-CMA-II diagnosis code = 29900, 29901, 29910	
Dementia		Chemical ID = 3923, 3750	ICD-10-AM diagnosis code = F00-F03 ICD-9-CMA-II diagnosis code = 290, 2941	
Eating disorder	Team type = 16		ICD-10-AM diagnosis code = F50 ICD-9-CMA-II diagnosis code = 3071, 30750,	

[10] PRIMHD is the Ministry of Health's national collection of activity and outcomes data. It stands for 'Programme for the Integration of Mental Health Data'.

Diagnosis class	PRIMHD[10]	Pharmaceuticals	NMDS – publicly funded hospital discharges	MSD incapacity
			30751, 30754, 30759	
Intellectual disability	Team type = 12		ICD-10-AM diagnosis code = F70 - F79 ICD-9-CMA-II diagnosis code= 317-319	Incapacity code = 008, 164
Mood		Chemical ID = 1069, 1437, 1438, 2466, 1824, 3753, 3901, 1125, 1955, 2285, 2301, 6009, 2636	ICD-10-AM diagnosis code= F30-F39 ICD-9-CMA-II diagnosis code = 296, 3004, 30113, 311	
Mood anxiety		Same codes as for Anxiety or Mood and/or Chemical ID = 2632, 1193, 3926, 1760, 2638, 3927, 1030, 1180, 3785	Same codes as for Anxiety or Mood	
Other/ Undetermined	All remaining records (where Activity type ≠ T09, T16, T17, T18, T19, T20 and/or Team type ≠ 03, 10, 11, 12, 16, 21, 23)	Not otherwise diagnosed in this table and Chemical ID = 1080, 1729, 1731, 2295, 1315, 1533, 1535, 1140, 1911, 1950, 1183, 1011, 3873, 1642	Not otherwise diagnosed in this table and any other the following: ICD-10-AM diagnosis code = F04-F09, F51-F53, F59, F63, F68, F69, F930-F932, F99 ICD-9-CMA-II diagnosis code = 2930-2940, 2948, 2949, 29911-29991, 30016, 30019, 30151, 3027, 30651, 3074, 30921, 310, 3123, 3130, 3131	Incapacity code = 009, 160, 163, 165
Personality			ICD-10-AM diagnosis code = F60-F62 ICD-9-CMA-II diagnosis code = 3010, 30110, 30111, 30112, 30120-30150, 30159, 3016-3019	
Potential mental health and addiction conditions		Chemical ID = 1389, 4037, 1059, 1379, 1876, 1190, 1316, 1397, 1730, 6007, 1283, 2298, 2530, 3803, 3898, 8792, 1273, 1865, 1956,		

Diagnosis class	PRIMHD[10]	Pharmaceuticals	NMDS – publicly funded hospital discharges	MSD incapacity
		2224, 2539, 3248, 1583, 1799, 3940, 4025, 3722, 3892, 3920, 3950, 1795, 1007, 1013, 1111, 1226, 1252, 1578, 1841, 2436, 3735, 3935, 1002, 1217, 2166, 2484		
Schizophrenia and related psychotic conditions	Activity type = T09	Chemical ID = 3884, 3878, 1078, 1532, 2820, 1732, 1990, 1994, 2255, 2260	ICD-10-AM diagnosis code = F20-F29 ICD-9-CMA-II diagnosis code = 2950-2959, 2970-2989	Incapacity code = 161, 162
Substance use	Team type = 03, 10, 11, 21, 23 and/or Activity type = T16, T17, T18, T19, T20	Chemical ID = 2367, 1432, 3793	ICD-10-AM diagnosis code = F10-F16, F18-F19, F55 ICD-9-CMA-II diagnosis code = 291, 292, 3030-3050, 3052-3059	Incapacity code = 006, 007, 170, 171, 172

The diagnosis available in the PRIMHD dataset has not been used as per advice received by the Ministry of Health. There are known data quality issues and the field is often left blank.

Many classification codes can be prescribed for multiple reasons. For example: drugs such as Risperidone (code 1011) that can be used for treatment of Autism, Anxiety or Schizophrenia. The "other/unknown" class includes those codes that are mental health and addiction related but could apply to more than one type of condition.

The pharmaceuticals dataset lacks enough information to distinguish why a drug was prescribed. Analysts are advised to avoid the potential mental health and addiction conditions category diagnostics class unless they have a specific purpose which requires their inclusion. This category is based on medications that can be used for treatment of mental health and addiction conditions, but also for other conditions. For example: Amitriptyline is used to treat both mental health conditions and sleep disorder – we are not classing sleep disorder as a mental health and addictions condition. Hence we cannot be confident that an individual experiences a mental health and addiction condition based on this class alone. They are included in the table above for completeness, but are excluded in our discussion below.

Overlap of individuals across mental health and addictions data

The total number of individuals in the available datasets who have accessed mental health and addiction services or treatments in 2014 is just under 700,000. This equates to 15.5% of New

Zealand's 2014 estimated resident population (4.51 million). Comparing this to the Te Hinengaro 2006 12-month prevalence figure of 20.7% we likely have not identified all individuals with a MHA condition.

Table 5 identifies the overlap of individuals identified in the five administration datasets as having accessed mental health and addiction services or treatments in 2014. Yes = individuals are present in the data source, No = individuals are not present in the data source. For example there were 301,323 individuals who were dispensed pharmaceuticals in 2014 for mental health and addiction related use and did not appear in the other four datasets and there were 8,301 individuals who had only MHA pharmaceuticals dispensed and used publicly funded secondary services recorded in PRIMHD.

Table 5 Overlap of individuals across the 5 key MHA related datasets available in the IDI, 2014

Pharms dispensed	Secondary PRIMHD	Hospital admissions	Medical certificates	Laboratory tests	Count of individuals[11]
Yes	No	No	No	No	301,323
No	Yes	No	No	No	98,859
No	No	Yes	No	No	93,996
No	No	No	Yes	No	32,976
No	Yes	No	Yes	No	19,572
No	Yes	Yes	No	No	16,149
Yes	Yes	No	No	No	8,301
Yes	No	Yes	No	No	7,425
No	Yes	Yes	Yes	No	5,748
Yes	No	No	Yes	No	3,021
Yes	Yes	No	Yes	No	1,809
Yes	Yes	Yes	No	No	1,440
No	No	Yes	Yes	No	1,422
No	No	No	No	Yes	636

[11] Health data released from the IDI using the mental health definition codes released by SIA is suppressed for counts <=20 individuals.

Pharms dispensed	Secondary PRIMHD	Hospital admissions	Medical certificates	Laboratory tests	Count of individuals[11]
No	Yes	No	No	Yes	576
Yes	Yes	Yes	Yes	No	528
No	Yes	Yes	No	Yes	279
No	Yes	No	Yes	Yes	180
Yes	No	Yes	Yes	No	123
No	Yes	Yes	Yes	Yes	114
No	No	No	Yes	Yes	57
Yes	No	No	No	Yes	48
No	No	Yes	No	Yes	48
Yes	Yes	No	No	Yes	42
Yes	Yes	Yes	No	Yes	21
Yes	Yes	No	Yes	Yes	21
Yes	Yes	Yes	Yes	Yes	<=20
Yes	No	No	Yes	Yes	<=20
Yes	No	Yes	No	Yes	<=20

These figures can be used as a sense check for analysts working on building their own mental health and addiction conditions dataset for analysis. Please note the exact numbers will vary based on the population definition, time period and IDI refresh. The population above includes only those individuals who are linked to the IDI spine, and are part of the estimated resident population as at July 2014.

What time periods can be analysed?

The administrative sources used to identify individuals who are accessing services or treatments are available over different time periods. The common overlap between these data sources is 2008 – 2017 and will improve as more data is added to the IDI.

While PRIMHD data is not considered complete until 2012[12], as it does not include completed data for 2008 and 2009, analysis can be performed from 2008. This is because the overlap between the other datasets and PRIMHD is high enough that most individuals will be captured from 2008.

The pharmaceuticals deemed to be related to treatment of mental health and addiction conditions are relevant from at least 2005 to the latest data available in the IDI.

[12] The Ministry of Health has the view that PRIMHD data is not complete until 1 July 2012.

Data quality and reliability

Data does not need to be perfect to be useful for more informed decision making

The data we have gathered on mental health and addiction conditions is by no means a complete picture, nor is it without quality issues. This work has been completed on the principle that data does not need to be perfect. Data quality and availability will improve in response to increased demand for insights. SIA is committed to helping increase the quality and availability of data – at the same time we are pragmatic about making the most of what is available.

While each dataset does have quality issues, the data is still suitable for analysis, so long as quality issues are considered when drawing conclusions from the data.

There are three key caveats that impact the scope of possible analyses:

1. Limited addiction related data,
2. Quality issues in available data sources,
3. Several data sources in the IDI were not suitable for use.

Analysing addiction is currently limited in the IDI

The Ministry of Health defines three types of addiction; alcohol and other drug (substance), problem gambling, and smoking.

Addiction events in our consolidated table are limited to one type of addiction: alcohol and other drug abuse (also referred to as substance use disorder). This is because service or treatment data does not include information about harmful or problem gambling, so harmful or problem gambling is excluded from our definition. Smoking is also excluded as it is not adequately captured in the available data.

Quality issues with the available data sources

During the course of this work, SIA staff developed an understanding of the collection processes and quality of the key data sources. These are described below.

PRIMHD data quality

PRIMHD is a Ministry of Health collection of national mental health and addiction condition information regarding service activity and outcomes for healthcare users. The data is collected from DHBs and NGOs who provide mental health and addiction services with government funding (Vote Health). PRIMHD also contains information on patients who use alcohol and drug counselling services funded by the government (Vote Health only).

PRIMHD was established in July 2008. Prior to PRIMHD, mental health data was collected in the Mental Health Information National Collection (MHINC) and stored in the Mental Health Data Warehouse (MHDW). The MHDW was started in July 2000. This data is now available in the IDI for researchers.

Increased NGO reporting will influence trends

Shifts or patterns in the data after 2008 may reflect the gradual inclusion of NGOs into the PRIMHD collection in addition to, or instead of, any trend in mental health service use or outcomes. Although NGO data is still incomplete, the Ministry of Health considers it complete enough for comparison across time from 1 July 2012 onwards.

Completeness of data for older people

Mental health and addiction services for the elderly are funded as mental health and addiction services in the Northern and Midland regions. However, in part of the Southern and Central regions they are funded as disability support services. PRIMHD mainly captures mental health and addiction services, and occasionally captures data on disability support services. This means data on healthcare users aged over 65 (including psychogeriatric services) is incomplete.

Some organisations are behind on their reporting

Some NGOs and DHBs have not reported for several years due to changing systems, organisational changes etc. which creates gaps in data. The Ministry of Health works with organisations if they change their reporting method to PRIMHD to ensure minimal issues with data completeness.

The diagnosis code in PRIMHD is not recommended for data analysis

The Ministry of Health has advised that for a number of reasons many clients will have no diagnosis recorded, and where they do these codes are often of low quality. NGOs are unable to submit diagnosis data, and the Ministry of Health does not require a diagnosis within the first three months of treatment.

This means that there is no diagnosis recorded for more than half of clients, particularly those who only received short term treatment. Many service organisations also submit a large number of non-specific diagnoses such as 'diagnosis deferred' or 'no specific diagnosis' during this time.

In addition to diagnosis coding issues, users need to keep in mind that clinicians treat symptoms and diagnoses. This means use of any diagnosis identified through data should be used with care.

This work has identified situations where clinicians may use certain diagnosis codes in order to ensure medication funding is available and cases where clinicians are hesitant to record a diagnosis due to the risk of patients not being able to access certain insurance. As with all data collected for administrative purposes, the collection of data in real world circumstances can often be different to what is recorded in data dictionaries or what is expected. A deep understanding of how diagnosis is captured is required, otherwise this variable is not recommended for data analysis.

Changes in activity codes may cause artificial variance and trends

Observed variance and trends may be a result of changes in how activity codes are assigned across service providers and time, for example, activity code changes have influenced the number of crisis contact services reported by some DHBs.

To assist with assigning activity codes, the Ministry of Health has published the Guide to PRIMHD Activity Collection and Use[13]. As this guide was published in early 2016, the effect of the guide won't be seen in PRIMHD until the publication of 2015/16 data. A high level description of each of the activity types can also be found within the PRIMHD Codeset[14].

Data is continually updated and revised

PRIMHD is a living data collection, which continues to be revised and updated as data reporting processes are improved. Previously published data may be liable to amendments. There were notable changes made to the coding of team types as part of the HISO review of the PRIMHD Code set. Team type data, extracted before 1/7/2014, should not be compared with the data presently in the IDI.

To function as a national collection, PRIMHD requires integration with a wide range of patient management systems across hundreds of unique service providers. As the services adjust to PRIMHD, it is expected that the quality of the data will improve.

Referral end date issues

There are known issues with the end dates of PRIMHD referrals. Referral end dates may not be an accurate measure for when a client ceased accessing services. The referral may no longer be active but lack an end date, or the end date may reflect an administrative close.

Inactive referrals are common. It is believed that up to 30% of all open referrals in PRIMHD at the end of 2016 had not had any activity in the last 15 months. The reasons for these are varied and not fully known, for example:

- clients can be enrolled with multiple services from the same provider, but all the activity is recorded against a single service.
- providers may no longer be providing any services (due to loss of contract or closure, etc.) but the referrals have been left open in PRIMHD.

The Ministry of Health manages this through data quality checks and helping organisations to do a bulk-closure of referrals. These are one cause of administrative discharges.

Administrative closures also arise when organisations change their reporting method to PRIMHD or their patient management system. They may close all current referrals and open up new ones for clients. Mergers or handovers between organisations often require new organisation codes in PRIMHD, which also requires mass discharge and re-referral for existing clients.

General data quality issues

The relevance of the issues described below depends on what the data is being used to analyse. General quality issues include:

[13] https://www.health.govt.nz/publication/guide-primhd-activity-collection-and-use

[14] https://www.health.govt.nz/nz-health-statistics/national-collections-and-surveys/collections/primhd-mental-health-data/primhd-standards

- PRIMHD has duplicate or overlapping records (for activity, referral, etc.) – this does not affect distinct client counts, but does mean that in some cases, level of service use/activity volumes will be overstated
- invalid code combinations used, e.g. PH (telephone contact) setting for a bed night, etc.
- Electro-convulsive therapy/seclusion reported by NGOs (who do not provide such services)
- incorrect reporting of bed nights, e.g. records that start and end on the same day giving a zero bed night count
- team type/activity type mismatch, e.g. residential teams providing contacts, community teams providing bed nights etc.
- client age/team type mismatch, e.g. maternal mental health team with data against the baby's NHI instead of the mother's
- invalid or default time portions reported in date/time fields – can mean the duration of activity records is unable to be accurately determined.

Pharmaceutical data quality

The Pharmaceutical Collection supports the management of pharmaceutical subsidies. It is jointly owned by the Ministry of Health and PHARMAC. The Pharmaceutical Collection contains claim and payment information from community pharmacists for subsidised dispensed medications from 2005.

Pharmaceutical data does not have complete coverage of prescriptions for the treatment of mental health and addiction conditions

The Pharmaceutical Collection covers pharmaceuticals dispensed from community pharmacies. Research has shown that pharmaceuticals prescribed are sometimes not dispensed, for reasons such as cost. More specifically, in New Zealand research last decade, 7% reported that they did not collect prescription drug/s at least once during the preceding 12 months because of cost and people who reported high and very high levels of psychological distress (20% and 33% respectively) were more likely not to collect prescriptions than those reporting low levels of psychological distress (4.2%) (Jatrana, Crampton, Carter, & Richardson, 2008).[15] It is expected many people with very high psychological distress will be captured through PRIMHD data.

Pharmaceutical data can inflate mental health and addiction access numbers

The pharmaceutical data, if used as is, may inflate the number of people who access mental health and addiction services. Some of the drugs can be used to treat both MHA and non-MHA conditions; for example, some anti-depressant category drugs can also be used to treat chronic pain, or menopausal symptoms. For this analysis, SIA has taken into assumption that drugs that can be related to mental health and addiction conditions are prescribed to treat MHA conditions. An effective way of handling this inflation is to exclude drugs under the "potential MH" diagnostic class.

[15] We quote percentage from the relevant tables in this report rather than the body text because the body text for prescriptions (chapter 8) often appears to incorrectly repeat percentages from the previous chapter that relate to visiting the doctor rather than prescriptions (chapter 7).

There are duplicates in the pharmaceutical data

The Pharmaceutical Collection is known to include duplicate records. The majority of such duplicates are explained by dispensing of the same drug, but of different weights. For example: a pharmacist may dispense 75g of a drug by dispensing a 50g tablet and a 25g tablet. This will appear as two entries in the table.

The other source of duplicates is due to the credit, resubmit and reversal process, which relates to the funding of medication. There are also resubmits/reclaims data duplicates in the table that are not exact duplicate rows, but differ in a single column value or two. The Ministry of Health was consulted on this issue however it was unable to be resolved.

National Minimum Dataset quality

The NMDS is a national collection of hospital discharge information, including clinical information for inpatients and day patients from all of the publicly funded hospitals and discharge data from about 50% of the privately funded hospitals. The public hospitals have been updating data in an electronic format since 1993 but the data has been made available from 1988.

Only publicly funded hospital events are available in the IDI

Not all privately funded hospitals in New Zealand submit their data to the Ministry of Health, and so only publicly funded hospital events have been considered in our table. The implication of this is that it will exclude those who privately fund their hospitalisations, and so undercount the total number of individuals receiving mental health and addiction treatment.

MSD Incapacity Dataset quality

The MSD Incapacity Dataset identifies people who receive Job Seeker Health Conditions and Disability (JS-HCD) or Supported Living Payment (SLP). This dataset records the reason for their incapacity to work using incapacity codes, which allows this analysis to identify individuals who have received a medical certificate application or renewal. In the IDI, the data is available from 1993.

The source of the medical certificate is unknown

The data does not capture where the medical certificate was issued. This means it is not possible to distinguish between primary or secondary care or public or private care using this data source.

MSD Incapacity Data may miss some people who access mental health and addiction services

Up to four reasons for being unable to work, or incapacity codes, can be recorded for each person. As these codes are prioritised, only the primary incapacity code was used to determine whether mental health was listed as the reason for incapacity. However a broader analysis should also consider the remaining incapacity codes.

Laboratory Claims Collection quality

The Laboratory Claims Collection contains claim and payment information for laboratory tests with data from 2003. It does not contain any information about tests conducted in hospital laboratories that have been processed by the General Transaction Processing System at the Ministry of Health. It also contains laboratory test information from Pegasus and Medlab South Independent

Practitioner Association providers. As at February 2011, this amounted to over 275 million rows of claim and payment data.

There are many duplicate records in laboratory claims data

There are duplicate records in the data with a difference only in the lab test number field ("MOH_LAB_TESTS_NBR"). It is unclear if each lab claim is valid so we're including them all, each with separate lab claims costs at this point.

There are many labs with zero cost attributed in the records. This is because of the change to bulk funding and estimated costs for an individual event are now available in the IDI. The field MOH_LAB_EST_EXCL_AMT is now on the MOH Laboratory Claims table in the IDI and represents the estimated cost for the test(s).

Several data sources in the IDI were not suitable for our purposes

Three data sources available in the IDI were not used in this analysis as they were not in a format that was able to be analysed to determine MHA service usage:

1. Primary healthcare enrolment data
2. Disability data
3. Privately funded hospital discharge data.

Primary healthcare enrolment data is excluded

The primary healthcare enrolment data records whether an individual was actively enrolled with a primary healthcare organisation (PHO), most commonly via their general practitioner (GP). The current GP visits dataset in the IDI is of limited utility since it only records registration with a GP and doesn't include the reason for visit or patient diagnosis.

The IDI currently has very limited information on primary health care services. The PHO enrolment data contains information about the demographics of the population cohort enrolled in a PHO, along with when they were enrolled, when they were last seen, and the practice type. There are no details available to show what the GP visit was for, or the outcome of the visit. The data is stored in quarterly snapshots, so it is also not possible to see the exact date a person visited a GP, it is only possible to tell if they had contact in the quarter. Hence this data currently has limited use to inform mental health and addiction conditions insights.

It is possible to construct a frequency of primary care contact history at the individual level using three data sources in the IDI:

- Primary Healthcare Organisation (PHO) - last contact date
- Laboratory Testing Claims – provider visit date
- General Medical Subsidy Claims – GP visit date

This method over-counts contact frequency by 5-9%, as some laboratory tests are undertaken on behalf of public hospitals. Having the Medical Registration Number of the doctor requesting the

test and National Provider List primary care provider practising spells allows a refinement that excludes the laboratory test events that occur in public hospitals.

Disability data (SOCRATES) is excluded

The National Needs Assessment and Service Coordination Information (SOCRATES) is used by Ministry-funded Needs Assessment and Service Coordination (NASC) agencies to record information about clients who are eligible for Disability Support Services (DSS). This data has been excluded as the outcome of an assessment (which tells you if a mental health condition and/or addiction were identified) is in a different table to the other assessment data, specifically the date of the assessment. Because there is no date available in the assessment outcome table, there is no way to do a one to one join with the other details from the assessment as a person can have multiple assessments, with only a single diagnosis. This means there is no way to tell when a mental health condition or addiction related outcome was identified.

Changes to the structure of data would mean it could be used to improve mental health and addiction conditions insights.

Privately funded hospital discharges data is excluded

This data has been excluded because, at the time of writing this report no records for discharges from privately funded hospital visits were available in the IDI. Hence there was no way to determine who received a diagnosis or when.

Privately funded hospital discharge information is held by the Ministry of Health for about half of all events and is now available in the IDI. It is yet to be included in the consolidated data, but could be done so following the same process as publicly funded discharges.

Strengthening MHA data in the IDI

Understanding who needs support, who receives services, and how well those services work for the people receiving them is crucial to effective sector decision-making. To support this, the data in the IDI could be strengthened in the following ways:

- improving data already in the IDI
- adding data already collected into the IDI

Examples of more detailed questions the IDI could answer if these improvements were made are provided in section “What are some of the questions the available IDI data can answer?”.

Improving data already in the IDI

Throughout this work, data dictionaries, data quality information, and advice on the use of data was sourced from many different places: directly from data providers, the Ministry of Health, Statistics NZ analysts, and the IDI researcher community, as well as agencies’ websites and the IDI wiki (documentation space).

Collecting this information and understanding how it relates to the use of data required a disproportionate amount of time and effort in the project.

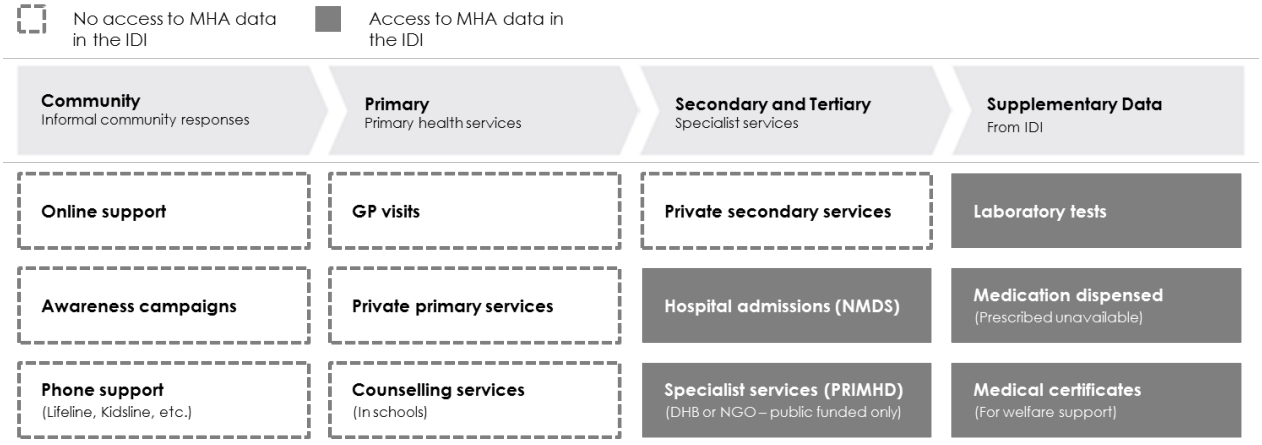
Detailed metadata, including data dictionaries, quality notes, changes to data over time, derived variables, and details about how the data was collected, should be standard information provided with any data source. This information is crucial for analysts to use data in the correct and accurate way, and should be available to all analysts.

Information on the cost of services and treatments also needs to be improved. Knowing how much a service or treatment costs is crucial to more informed investment decisions. The quality of cost information of all social services available in the IDI (not just mental health condition and addiction related data sources) varies extremely; from no costs, to estimated costs, to actual costs.

Using actual and estimated costs, SIA was able to map 60% of government spending on social services to individuals (Social Investment Unit, 2017). Cost coverage has since increased, and SIA is seeking to further improve the coverage. The higher this percentage the more accurate cost analyses will be possible.

Adding data from administrative data sources into the IDI

Figure 4: Administrative data available in the IDI of individuals who have accessed mental health and addiction services or treatments



Many data sources which could improve mental health and addiction conditions insights are not available in the IDI. These include:

- Primary health care
  - PHO data relating to patient visits and services provided by GP practices (currently only enrolment information is available in the IDI)
- Secondary health care
  - MHINC (2000-2007). This is the dataset that precedes PRIMHD. It is held by the Ministry of Health
  - interRAI Home Care Assessment captures information about elderly people's function and quality of life by assessing needs, strengths, and preferences, and is held by the National interRAI Data Analysis and Reporting Centre
  - Health of Nations Outcome Measure (HONOS) data in PRIMHD
  - Alcohol and Drug Outcome Measure (ADOM) data in PRIMHD
  - Supplementary Consumer Record (SCR) data in PRIMHD
- Non-health care
  - 1X and 1M police calls, which relate to calls made to the NZ Police for attempted suicide and mental health respectively.

Of these, particular emphasis should be given to primary health care information held by PHOs. There are currently 32 PHOs that ensure the provision of essential primary health care services mostly through GPs.

SIA research has shown that the addition of more PHO data, relating to services and treatments provided by GP practices into the IDI would provide a clearer picture of the full treatment pathway for mental health conditions or addiction.

- It can help us quantify people who fail to collect MHA medications prescribed to them by a GP or any other primary health care specialist.
- The subset of the population receiving non-pharmaceutical based MHA treatment through primary care can be identified.
- PHO data also provides a more complete picture of an individual's treatment history, to supplement information already available through the IDI.

Making PHO data available in the IDI for MHA related research would thereby provide fine-grain details regarding an individual's pathway into the MHA system and also plausible conditions on their treatments and services. It would be highly beneficial in terms of segmentation of service use and uptake of services.

Additional survey data would also assist

Several surveys exist that measure mental wellness which could improve mental health and addiction conditions insights, but are not currently available in the IDI. These include:

- The New Zealand Health Survey (NZHS), 2011/12 – 2017/18 (annual)[16]
- Dunedin Multidisciplinary Health & Development Study, 2017-2019 (Mental Health, Phase 38, up to the age of 45 Years)
- Te Rau Hinengaro: The New Zealand Mental Health Survey, 2006.

Adding the NZHS to the IDI would not only provide a more up to date source of information compared with SoFIE (2013/14 compared with 2008/09), it also contains more measures of substance use such as the updated delivery of the AUDIT questionnaire in 2015/16, which has been shown to increase the proportion of respondents reporting hazardous drinking. This data is currently not available in the IDI due to the sensitivity of respondent consent.

These additional questions about addiction could be used to provide insights about addiction compared to service use, as we did for mental health conditions (K-10) in Section "What is the K-10?". Currently SoFIE does not provide sufficient questions about addiction to perform this analysis.

Neither the NZHS nor Te Rau Hinengaro captures population level information on psychosis. This is a data gap both in and outside of the IDI.

[16] NZHS can be accessed as a confidentialised unit record file (CURF). To access CURFs, the research and the researchers need to meet certain assessment criteria and obtain approval of access.

What are some of the questions the available IDI data can answer?

Making the best possible use of data and analytics, as well as other evidence, helps the social sector to get the right help to the people who need it at the right time. For MHA, particular concerns revolve around understanding unmet need.

Table 6: Examples of questions the IDI can and cannot help answer relating to mental health and addictions. This table does not represent all possible questions, but provides an illustration of the types of questions that can and cannot be explored in the IDI now, or with additional data

High level question	The IDI can currently be used to help answer these questions	The IDI could be used to answer these questions if existing data is added into the IDI or made available to more users in the IDI	The IDI could be used to help answer these questions if data is gathered and made available in the IDI	It isn't appropriate to use the IDI to help answer these questions[17]
Who needs support?	What macro-level groups appear to need mental health services, but are not receiving them? For example: Are Māori living in Gisborne under-serviced for mental health services?	What macro-level groups appear to need addiction services, but are not receiving them? For example: Are Māori living in Gisborne under-serviced for addiction services?	Which (de-identified) individuals need support?	Service delivery support – which individuals should we give services to? For example: Does person A or B need a service? What causes mental health conditions or substance use disorder?
When is the best time to receive support?	What are common life courses for people who use MHA services or treatment (in the IDI)? For example: What are the government 'touch points' throughout a person's life,	What are common life courses for people who use any MHA services or treatment? For example: What are all of the government 'touch points' throughout a	Does a service or treatment delivered early in life have a better impact than a service or treatment delivered later? (Question also repeated below as it is an impact question)	What is the medical recommendation for services or treatments?

[17] Data at the level of granularity required to answer these questions may not reflect the actual status of service/need.

High level question	The IDI can currently be used to help answer these questions	The IDI could be used to answer these questions if existing data is added into the IDI or made available to more users in the IDI	The IDI could be used to help answer these questions if data is gathered and made available in the IDI	It isn't appropriate to use the IDI to help answer these questions[17]
	and where would be best to intervene? Do people who are referred to specialist services from services such as primary care or school have better outcomes compared to people who are referred from crisis services such as Justice or ED?	person's life, and where would be best to intervene?		
What services or treatments will be most effective for different groups?	Results from any of the impact evaluation question below, but all results will be at least four years old due to IDI lag time.	Results from any of the impact evaluation question below, but all results will be at least four years old due to IDI lag time.		How well does a new service work (before it is delivered)? For example: How well do e-services work for people?
Who accesses services or treatments?	Who received mental health or addiction services or treatments (available in the IDI) and what are their characteristics? For example: How old were people when they received services (in the IDI) for the	Who received any mental health or addiction service or treatment and what are their characteristics? For example: How many people received services or treatments (an accurate and complete count)?		What individuals are receiving services today?

High level question	The IDI can currently be used to help answer these questions	The IDI could be used to answer these questions if existing data is added into the IDI or made available to more users in the IDI	The IDI could be used to help answer these questions if data is gathered and made available in the IDI	It isn't appropriate to use the IDI to help answer these questions[17]
	first time? Are services (in the IDI) delivered equitably across socio-economic status? What experiences do people have before or after an event or before or after receiving an MHA service or treatment? For example: What government interactions do people have before they self-harm? What mental health services or treatments do people receive before and after offending?	How old were people when they received services for the first time? Are services delivered equitably across socio-economic status?		
How well do services or treatments work for the people accessing them (impact evaluation)?	What is the impact of services and treatments (in the IDI) on non-mental wellness outcomes? For example: Do services delivered by NGOs result in more employment for the	What is the impact of any services and treatments on non-mental wellness outcomes? For example: Do services delivered by GPs result in more employment for the	What is the impact of any service or treatments on mental wellness outcomes? NB: services are designed to support mental wellness. This is the better	Timely and operational evaluation of services, where timely insights can feedback into service delivery. For example: Should we give person A service 1 or

High level question	The IDI can currently be used to help answer these questions	The IDI could be used to answer these questions if existing data is added into the IDI or made available to more users in the IDI	The IDI could be used to help answer these questions if data is gathered and made available in the IDI	It isn't appropriate to use the IDI to help answer these questions[17]
	people receiving them, compared with services delivered by DHBs?	people receiving them, compared with services delivered by secondary services?	measure (compared with other outcomes) to measure service effectiveness. Does a service or treatment delivered early in life have a better impact than a service or treatment delivered later?	2? Person B started receiving service 2 three months ago, is it working? Should we make changes to the delivery?

Appendix

Can self-harm and suicide be identified in available data?

This can be done using the data sources and codes in the table below.

Table 7: Data and codes available in the IDI to identify self-harm, suicide attempts or suicide

Data source	Self-harm, suicide attempt codes, or suicide codes	Cause of death codes
ACC injury claims	acc_cla_wilful_self_inflicted_status_text = 'CONFIRMED'	
CYF abuse events	cyf_abe_source_uk_var2_text in ('SHM', 'SHS', 'SUC')	
Hospital admissions (NMDS)	External causes (diagnosis type = 'E') + ICD10 code between X60-X84 (intentional self-harm). Also include ACC claim is NULL as ACC events will be double counted if you are using ACC data as well	
Mortality data (Ministry of Health)		moh_mor_icd_d_code in (X60-X84) (Intentional self-harm)
CARD 111 1x calls (Police Integrated communication centre database and National Intelligence Application data)	nia_links_latest_inc_off_code = '1X' Note: Only caller identified data is integrated into the NIA data.	

Note ACC injury data and hospital admission data both record if a person is deceased. However the mortality data is the only source which uses a standard process to code mortality. This is the recommended source and is comparable internationally.

The mortality data has a time lag getting into the IDI, so current data is always a few years behind.

This data could be improved using self-harm and suicide calls made to the New Zealand Police, known as 1X and 1M calls. These codes are made by frontline police staff using their judgement about the nature of the call they are receiving. However, only a preliminary version of this data is available in the IDI.

Pharmaceuticals used in mental health and addiction treatment

Table 8 shows the chemical names and type.

Table 8 Chemical ID names and types

Chemical ID	Chemical name	Type
1011	Risperidone	Antipsychotics
1030	Sertraline hydrochloride	Antidepressants
1069	Amoxapine	Antidepressants
1078	Clozapine	Antipsychotics
1080	Amylobarbitone sodium	Sedatives and hypnotics
1125	Nefazodone	Antidepressants
1140	Olanzapine	Antipsychotics
1166	Bromazepam	Anxiolytics
1180	Venlafaxine	Antidepressants
1183	Quetiapine	Antipsychotics
1183	Quetiapine fumarate	Antipsychotics
1193	Citalopram hydrobromide	Antidepressants
1315	Clomipramine hydrochloride	Antidepressants
1432	Disulfiram	Treatments for substance dependence
1437	Dothiepin hydrochloride	Antidepressants
1438	Doxepin hydrochloride	Antidepressants
1532	Flupenthixol decanoate	Antipsychotics
1533	Fluphenazine decanoate	Antipsychotics
1535	Fluphenazine hydrochloride	Antipsychotics
1642	Imipramine hydrochloride	Antidepressants
1729	Loprazolam mesylate	Sedatives and hypnotics
1731	Lormetazepam	Sedatives and hypnotics
1732	Loxapine succinate	Antipsychotics
1760	Maprotiline hydrochloride	Antidepressants
1780	Meprobamate	Anxiolytics

Chemical ID	Chemical name	Type
1809	Methylphenidate hydrochloride	Stimulants/ADHD treatments
1824	Mianserin hydrochloride	Antidepressants
1911	Oxazepam	Anxiolytics
1950	Pericyazine	Antipsychotics
1955	Phenelzine sulphate	Antidepressants
1990	Pimozide	Antipsychotics
1994	Pipothiazine palmitate	Antipsychotics
2255	Thioridazine hydrochloride	Antipsychotics
2260	Thiothixene	Antipsychotics
2285	Tranylcypromine sulphate	Antidepressants
2295	Triazolam	Sedatives and hypnotics
2301	Trimipramine maleate	Antidepressants
2367	Calcium carbimide	Stimulants/ADHD treatments
2466	Lithium carbonate	Antipsychotics
2632	Alprazolam	Anxiolytics
2636	Fluoxetine hydrochloride	Antidepressants
2638	Moclobemide	Antidepressants
2820	Fluspirilene	Antipsychotics
3750	Rivastigmine	Stimulants/ADHD treatments
3753	Mirtazapine	Antidepressants
3785	Venlafaxine	Antidepressants
3793	Naltrexone hydrochloride	Treatments for substance dependence
3873	Ziprasidone	Antipsychotics
3878	Aripiprazole	Antipsychotics
3880	Methylphenidate hydrochloride extended-release	Stimulants/ADHD treatments
3884	Amisulpride	Antipsychotics
3887	Atomoxetine	Stimulants/ADHD treatments
3901	Mirtazapine	Antidepressants

Chemical ID	Chemical name	Type
3923	Donepezil hydrochloride	Treatments for dementia
3926	Escitalopram	Antidepressants
3927	Sertraline	Antidepressants
6006	Buspirone hydrochloride	Anxiolytics
6009	Paroxetine hydrochloride	Antidepressants
1002	Lamotrigine	Potential MHA
1007	Sulpiride	Potential MHA
1013	Guanethidine sulphate	Potential MHA
1059	Amitriptyline	Potential MHA
1111	Dexfenfluramine	Potential MHA
1190	Citalopram hydrobromide (Celapram)	Potential MHA
1217	Carbamazepine	Potential MHA
1226	Zuclopenthixol dihydrochloride	Potential MHA
1252	Naltrexone hydrochloride	Potential MHA
1273	Chlormethiazole edisylate	Potential MHA
1283	Chlorpromazine hydrochloride	Potential MHA
1316	Clonazepam	Potential MHA
1379	Desipramine hydrochloride	Potential MHA
1389	Dexamfetamine sulfate	Potential MHA
1397	Diazepam	Potential MHA
1578	Glycopyrronium bromide	Potential MHA
1583	Haloperidol	Potential MHA
1730	Lorazepam	Potential MHA
1795	Methadone hydrochloride	Potential MHA
1799	Levomepromazine maleate	Potential MHA
1841	Naloxone hydrochloride	Potential MHA
1865	Nitrazepam	Potential MHA
1876	Nortriptyline hydrochloride	Potential MHA

Chemical ID	Chemical name	Type
1956	Phenobarbitone sodium	Potential MHA
2166	Sodium valproate	Potential MHA
2224	Temazepam	Potential MHA
2298	Trifluoperazine hydrochloride	Potential MHA
2436	Flunitrazepam	Potential MHA
2484	Zopiclone	Potential MHA
2530	Haloperidol decanoate	Potential MHA
2539	Midazolam	Potential MHA
3248	Chloral hydrate	Potential MHA
3722	Nicotine	Potential MHA
3735	Melatonin	Potential MHA
3803	Zuclopenthixol decanoate	Potential MHA
3892	Bupropion hydrochloride	Potential MHA
3898	Zuclopenthixol hydrochloride	Potential MHA
3920	Varenicline tartrate	Potential MHA
3935	Modafinil	Potential MHA
3940	Olanzapine pamoate monohydrate	Potential MHA
3950	Buprenorphine with naloxone	Potential MHA
4025	Paliperidone	Potential MHA
4037	Rivastigmine	Potential MHA
6007	Chlordiazepoxide hydrochloride	Potential MHA
8792	Droperidol	Potential MHA

Subsets of the populations for specific purposes

The purpose of the consolidated mental health and addiction table generated using the code available in GitHub is to identity all MHA-related events in the IDI and assemble them into a consistent structure. The main intention behind this is to help establish common ground for MHA based research. Researchers interested in service use or intensity of use will wish to use a subset of the full table. For example, when studying service use a person who was referred to a secondary mental health and addiction service (captured in PRIMHD), but did not attend an appointment (captured in the data by DNA = did not attend) should be excluded.

Both the subsets described below were developed with the Ministry of Justice (Lunt, L. W. (2017)) and the Ministry of Health. They are subsets of the main consolidated table defined in the body of this document to ensure all analysis starts from a consistent point.

Identifying service use

The code that SIA developed brings together data on MHA conditions using the five data sources mentioned in Table 9 Possible subsets of the population who have accessed services or treatments. To consider different combination of events where people received mental health and addiction services we recommend modifying the consolidated table according to the description of the datasets mentioned below.

Table 9 Possible subsets of the population who have accessed services or treatments

IDI dataset	Mental health and addiction service or treatment access	Mental health and addiction service or treatment use	Notes
PRIMHD	Any record in PRIMHD (except for activity types T47 and T49)	Face-to-face PRIMHD activities, excluding where the client was not present[18]	To have measure of actual services used (as a measure of mental health intervention) then excluding activities when the client is not present, did not attend, or when the activity was not face-to-face – codes based on advice from MoH. T47 and T49 codes are assigned to a family member's NHI number – therefore excluded as the service is not about their own mental illness
Pharmaceutical dispensing data	Pharmaceuticals deemed to be MHA-related (excluding potential MHA category)	As per SIA definition	See also the work by the Ministry of Justice on mental health service use in 2015 (Lunt, L. W. (2017))

[18] Activity_type_code not T08 (care/liaison co-ordination contacts), T32 (contact with family/whānau, consumer not present), T33 (seclusion), T35 (did not attend), T37 (on leave), T47 (support for family/whānau), T49 (support for children of parents with mental illness and addictions). And activity_setting_code not SM (text messaging), PH (telephone), WR (written correspondent), OM (other social media / e-therapy).

IDI dataset	Mental health and addiction service or treatment access	Mental health and addiction service or treatment use	Notes
NMDS (publicly funded discharges only)	Any hospitalisation event with an associated MHA diagnosis	As per SIA definition with the exception that the MHA diagnosis needs to be primary diagnosis	Amendment to the service use definition since unless the MHA diagnosis is the primary diagnosis, then the hospitalisation may not have been related to MHA, but rather staff were aware of the patient's MHA
MSD Incapacity data	Any record with an MHA-related medical certificate for benefit support	Not included	Some benefit types (Jobseeker Support) have to regularly see a GP for medical certification, while other benefits (Supported Living Payment) is every 2 years or isn't necessary. While the former group are receiving mental health services from their GP, it is likely this group will already have been included through pharmaceutical dispensing
Laboratory claims (LAB) collection	More than 2 lab tests for lithium within a 4-month period	Not included	Having a lab test is not a mental health treatment or service. Further previous advice from MoH to exclude this since: (1) Results of the tests are not in the data, hence not conclusive being MHA-related; (2) People would already be picked up with the pharmaceutical dispensing data if they have lithium-based prescription drugs.

Subset for level of service use

Researchers wishing to identify people with a high use of MHA services could modify the consolidated table according to the table below.

Table 10 A set of data definitions to create different levels of service use

Category	Derivation of level of mental health and addiction service use
High use	Include people who meet one or more of the severe/long term/hospitalisation criteria as follows: ▪ ***Severe:*** Individual had face-to-face PRIMHD activity started in the 12 months either side of the chosen point in time. Activities reflecting service use for severe mental health and addictions problems include any of the following activity type codes: T02, T03, T04, T11, T12, T13, T14, T16, T20, T21, T25, T26, T27, T28, T29, T48 and/or face-to-face activity with team type 01 (inpatient) or 05 (forensic) ▪ ***Long term:*** In the 12 months either side of the chosen point in time: » Aged 10 to 16 years: Been provided new or on-going relevant** PRIMHD mental health activities in FOUR or more consecutive quarters intersecting at some point in 2-year period » Aged 17 years or more: Been provided new or on-going relevant** PRIMHD mental health activities in all EIGHT quarters of the 2-year period[19] **Notes:** a) Age is at the time of the chosen point in time b) Quarters are defined as 3-month periods counting back from the end of the 2-year period (specifically 91.325 days ignoring leap years) c) Relevant mental health activities include face-to-face activities except for family/whānau activities where these are targeted to the family/whānau rather than the patient (e.g. T47, support for family/whānau, T 49, support for children of parents with mental illness and addictions): activity type codes T08, T32, T33, T35, T37, T47 and T49 are excluded, as are activity setting codes SM, PH, WR, and OM ▪ ***Hospitalisation:*** with primary diagnosis mental health related (as per broader discharge diagnosis codes where diagnosis type = 'A') within 12 months either side of the chosen point in time.

[19] Please note that the activities do not need to be part of the same referral and all referrals for the person should be considered.

Estimating self-reported need in the population

Complementing MHA administrative data with surveys

Surveys are an integral dataset for understanding the occurrence of specific conditions in the population. The prevalence of mental health conditions was mapped by the seminal Te Rau Hinengaro: New Zealand Mental Health Survey[20]. The findings help support a deep understanding of mental health need and demand in the population, and also reveal some patterns when the need and demand are graphed against deprivation deciles.

Surveys like SoFIE and the NZGSS screen for mental health need using short-form questions on mental well-being, like the K-10 or SF-12.

The provision of MHA services or treatments by DHBs or NGOs is an effort to address the identified need in the population. But there can be a number of challenges which prevent equitable uptake of these services.

IDI data on service/treatments reflects the meeting of MHA needs and service uptake. This data can complement surveys in two key ways:

- Survey estimates of prevalence can be compared to uptake of services. *Are services reaching those in need?*
- Administrative touch points of surveyed individuals can be compared against their responses. *Do self-reports of need match service uptake?*

What is SoFIE?

SoFIE is a longitudinal panel design survey which measures how family, income and employment change over time. Data was collected each year from 2002 – 2010, which are referred to as survey waves.

SoFIE IDI weighting methodology and representativeness of the population

To produce population estimates from survey data it is necessary to use weights. This appendix discusses the adjustment SIA did to the SoFIE weights and how representative the SoFIE population is of the New Zealand adult population. This appendix also covers how similar results are between SoFIE and NZHS.

The group of people analysed are a subset of those in the SoFIE survey that we have enough information about to construct a measure of distress and to follow their life course

To estimate the number of people with mental health distress we chose to look at only those who answered the K-10 questions. The population is constructed as follows. Take the people who:

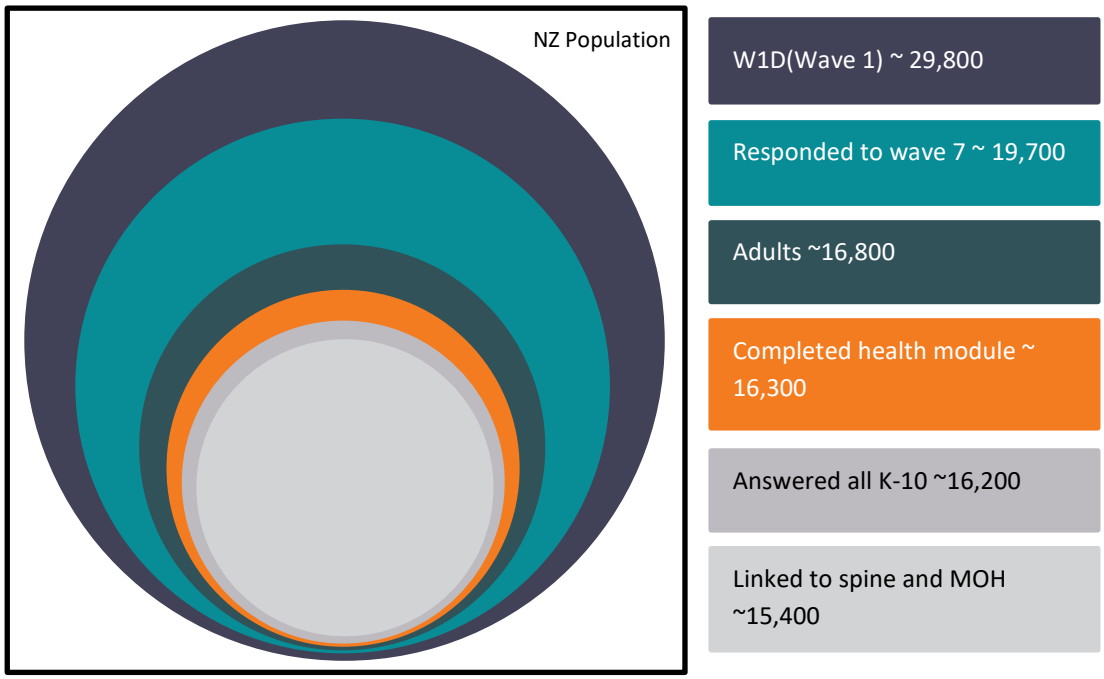
- provided a complete and eligible response in wave 1 (W1D)
- responded in wave 7
- are adults (children do not complete the health module in SoFIE)
- completed the health module

[20] https://www.health.govt.nz/publication/te-rau-hinengaro-new-zealand-mental-health-survey

- answered all K-10 questions
- are linked to the IDI spine and have a Ministry of Health ID (NHI number).

Note that this methodology is similar to other mental health work by Richardson, Carter, & Hayward (2008). The main difference is at step 5 where they allow for a maximum of one missing K-10 question which was then imputed. Given that there were so few people who didn't answer all the questions and the extra work to find a suitable imputation method for the survey, the decision was made not to follow their approach.

Figure 5: Population covered under the SoFIE survey and our subset of the population for mental health and addiction conditions along with number of individuals in the survey at each point (not drawn to scale)



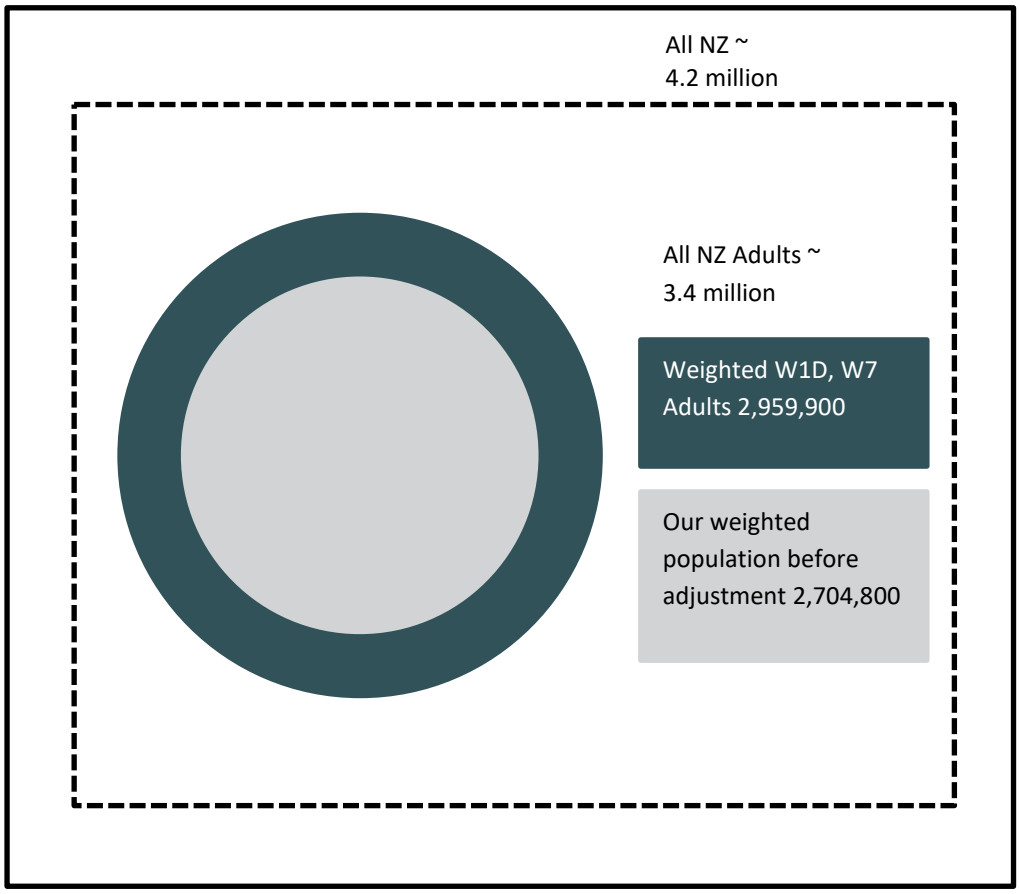
It is necessary to adjust the SoFIE survey weights to make inferences about the population

The SoFIE survey only provides data on the surveyed individuals but inferences can be made about the population. The SoFIE datasets provide a series of longitudinal weights that adjust for sampling bias, non-response and drop out over the waves. These weights are available in a SoFIE dataset in the IDI.

The available weights need to be adjusted to account for both records that do not link to the spine and for individuals who do not respond to the SoFIE health module, which includes the K-10 score.

The target population for the SoFIE survey is the resident population of New Zealand (adults and children) that live in private dwellings. Each person in the sample is weighted so that the sum of the sample weights is equal to the target population. We readjust the sample weights using cell adjustment so that our adjusted sample weights (the inner grey circle in Figure 6: Weighted sums of the adjusted population compared to the overall NZ population (not drawn to scale)) sum to the weights of the W1D, wave 7 adults (dark green circle in Figure 6: Weighted sums of the adjusted population compared to the overall NZ population (not drawn to scale)). This is done using region, age, gender and ethnicity to create a rate up factor to adjust the original sample weights by.

Figure 6: Weighted sums of the adjusted population compared to the overall NZ population (not drawn to scale)



It is important to note that the adjusted longitudinal weights will not sum to the size of the New Zealand adult population because the subset of first wave responders (W1D) in wave 7 excludes those in non-private dwellings and those who are not usually resident. Following the re-weighting process, the sum of the weights for the target population is 2,959,900.

The implication of this is we have no sight of mental health need for children and we have not covered all adults.

SoFIE health data was last collected in 2008-2009

SoFIE data used in this report is almost ten years old. Carter et al (2011) cites research that looked over a 10-year period and found "relatively stable rates of self-rated health across the 10 years" (K. N. Carter, Shaw, Collings, Blakely, & Richardson, 2011). This report concludes SoFIE is still a useful data source to use, in the absence of any other data sources to identify people who experience mental illness.

The SoFIE sign out rules differ to administrative data sign out rules

Any analysis must have un-weighted counts that exceed a frequency of 1 and weighted counts should be greater than or equal to 1000. All cell counts must then be rounded to the nearest 100. This will have an impact on how far you can drill down into the data.

How does SoFIE compare to the New Zealand Health Survey?

The NZHS is considered a robust and representative survey of New Zealanders' health needs. To ensure analysis in SoFIE is also robust and representative this section compares SoFIE to the NZHS (Statistics New Zealand, 2008).

In 2011, Carter and Shaw compared SoFIE and the NZHS and noted sources of biases that could produce differences in results including time bias and survey design differences. These included:

- **the population surveyed:** where the SoFIE population only includes those in private dwellings while the NZHS has respondents in private and non-private dwellings
- **survey design differences:** NZHS oversamples the minority ethnicity to get reliable estimates for them.[21] SoFIE has been stratified on socioeconomic variables. However, if weighted correctly, the different designs should not cause substantially different estimates.
- **survey length and question location:** The NZHS has the SF-36/SF-12 in the middle while SoFIE is considerably longer and has SF-36 towards the end and the health module was only available in some waves (2004, 2006 and 2008). This might result in SoFIE respondents being fatigued or conditioned in their responses. However, recent research (Carter et al, 2011) from the Dunedin Multidisciplinary Health and Development Study note that this was not the case in their experience of using repeated measures.

Carter et al. (2011) found the surveys share similar demographic characteristics which are described in more detail below.

Despite different survey designs, research suggests that SoFIE and NZHS have similar distributions for distress and demographic variables

New Zealand research indicates that although the SoFIE survey population is not perfectly representative of the New Zealand adult population, it is broadly representative in terms of what was most important for our research.

Carter et al (2011) reports that "SoFIE respondents tended to rate their health and health related quality of life better than participants of the 2006/07 NZHS. However, despite this there was no difference in the distribution of psychological distress between the two surveys. The relative ratio of rates to the overall average was similar between the two surveys, for social groups, ethnicity and area deprivation. So although health status was better in the SoFIE population, the patterns of differences among respondents in SoFIE and the 2006/07 NZHS were similar."

What data can be linked to SoFIE to increase information to analyse?

As SoFIE is linked to the IDI, survey data can be supplemented with administrative data to provide more information to analyse. We have shown an example of this using the mental health and addiction administrative services data. If an analyst would like to add additional administrative data it is recommended to only use characteristics, associations, and outcomes.

[21] This is done using a dual sample frame where Māori have a higher chance of selection from the electoral roll and also using a geographical sample frame where the selection probabilities have been modified so that areas with a higher density of ethnic minorities have a higher chance of selection.

SoFIE does not provide accurate estimates of social sector costs

Costs of social sector services are skewed throughout the population because all social sector services are not accessed by everyone in the population.

SoFIE is not designed to make estimates of this skewed variable, making estimates unreliable.

Can SoFIE be used to analyse mental wellness through time?

SoFIE can be used to analyse mental wellness through time, using data collected yearly between 2002 and 2010. The health module was included in the survey in 2004, 2006 and 2008 in waves 3, 5 and 7.

Table 11 SoFIE survey waves and health module availability

Wave	Year	Health module
Wave 1	2002-2003	
Wave 2	2003-2004	
Wave 3	2004-2005	Available
Wave 4	2005-2006	
Wave 5	2006-2007	Available
Wave 6	2007-2008	
Wave 7	2008-2009	Available
Wave 8	2009-2010	

What is the K-10?

The K-10 is a 10 item Likert scale for measuring psychological distress

It is a widely used self-reported measure that identifies those in need of further assessment for anxiety and depression. This measure was designed for use in the general population to detect high-prevalence mental health disorders (Deady, 2009). It asks questions about emotions a person has experienced in the last four weeks, and asks them to rate how often they experience each particular emotion on the following five point scale:

- none of the time
- a little of the time
- some of the time
- most of the time
- all of the time.

The full set of K-10 questions and their corresponding column names in the IDI are listed in Table 11.

Table 12 K-10 psychological distress scale screening questions

#	Question	IDI column name
1	How much of the time during the past 4 weeks did you feel tired out for no good reason?	Sofie_hlt_felt_tired_no_rsn_code
2	How much of the time during the past 4 weeks did you feel nervous?	Sofie_hlt_felt_nervous_code
3	How much of the time during the past 4 weeks did you feel so nervous that nothing could calm you down?	Sofie_hlt_felt_nervous_extreme_code
4	How much of the time during the past 4 weeks did you feel hopeless?	Sofie_hlt_felt_hopeless_code
5	How much of the time during the past 4 weeks did you feel restless or fidgety?	Sofie_hlt_felt_fidgety_code
6	How much of the time during the past 4 weeks did you feel so restless that you could not sit still?	Sofie_hlt_felt_fidgety_extreme_code
7	How much of the time during the past 4 weeks did you feel depressed?	Sofie_hlt_felt_depress_code
8	How much of the time during the past 4 weeks did you feel that everything was an effort?	Sofie_hlt_felt_depress_extreme_code
9	How much of the time during the past 4 weeks did you feel so sad that nothing could cheer you up?	Sofie_hlt_felt_effortful_code
10	How much of the time during the past 4 weeks did you feel worthless?	Sofie_hlt_felt_worthless_code

Questions 3, 6 and 8 are only asked if the preceding question does not have a response of “none of the time”. For example if someone says that in the last four weeks they have felt nervous none of the time for question 2, the interviewer would not ask “how much of the time during the past 4 weeks did you feel so nervous that nothing could calm you down?” i.e. question 3.

The K-10 scale runs from 10-50, with lower scores indicating wellness and higher scores indicating depression or anxiety. The tables within the IDI contain only the codified item responses. It is necessary to convert them into a K-10 score.

Since K-10 is contained in a survey it is necessary to group K-10 categories so that findings are robust

Analysing small numbers in surveys is problematic in two ways:

1. For the SoFIE survey which contains the K-10 responses, un-weighted cell counts must exceed 1 and weighted cell counts must be greater than or equal to 1000 to be signed out of the IDI
2. A large number of observations are needed to obtain robust estimates.

Table 13 Total number of people (un-weighted and weighted) under each K-10 bands

K-10 band (2 levels)	K-10 band (4 levels)	Frequency (SoFIE survey)	Frequency (population)	Standard deviation	Percentage
Low K-10	Low (10-15)	13,500	2,420,800	14,500	81.8
Low K-10	Mild (16-21)	2,000	370,300	9,300	12.51
High K-10	Moderate (22-29)	700	129,400	6,100	4.37
High K-10	Severe (30-50)	200	39,000	3,200	1.32

To ensure that there are enough observations to do cross-tabulations that can be signed out of the IDI, we have grouped the moderate and severe groups together to form a high K-10 group (Table 13).

- We have defined 'low K-10' as K-10 score <= 21.
- We have defined 'high K-10' as K-10 score >= 22.

How well does K-10 capture all mental health conditions?

The K-10 does an adequate job of capturing all mental health conditions

New Zealand research (Oakley Browne et al, 2010) has found the K-10 to be effective in discriminating between Composite International Diagnostic Interview (CIDI) 3.0 cases and non-cases for anxiety disorder, mood disorders and any study disorder. The same research also found that K-10 for the past month is an effective predictor of mental illness in the past 12 months.

Because the K-10 is used to indicate high prevalence mental disorders such as mood disorders, depression and anxiety, there is a risk that the K-10 will not capture other, lower prevalence mental health issues such as psychosis as reliably. Although this underestimation is expected to be low because of the low prevalence of these disorders (Andrews & Slade, 2001), this group of the population has the highest costs to social services relative to numbers and faces the largest inequities in health and social outcomes (Morgan et al., 2017). So it is very important we find a way of identifying this group and looking at their needs by adding additional data to the IDI (see Section Adding data from administrative data sources into the IDI).

Although K-10 does not address addiction issues, other studies that show a strong overlap between mental health and addiction conditions in New Zealand (The Mental Health Commission, 2008) suggest that K-10 will capture a certain proportion of people with addiction issues due to high comorbidity.

"Te Rau Hinengaro and other studies confirm that co-existing mental health and addiction conditions problems are the expectation rather than the exception throughout the New Zealand mental health and addiction conditions treatment system including in primary care."

The Mental Health Commission, 2018

How does K-10 compare to other measures used to understand mental wellness?

An alternative measure of mental wellness in the IDI is the SF-36

The Short Form Health Survey contains questions on eight different health scale dimensions including physical functioning, bodily pain, vitality, social functioning and mental health. It is a 0-100 scale designed to measure perceived disability; the higher the score the better the perceived health.

Figure 7: Graphical representation of the SF-36 self-reported health scale

Complete perceived disability — No perceived disability

The K-10 is easier for analysts to work with

The scoring of SF-36 is complex. More experienced users such as the researchers from the University of Otago have scored the SF-36 (Richardson et al, 2008). There is little publicly available information on scoring the SF-36 using the Medical Outcomes Study standard, making it difficult for new users to create score code. The K-10 is easier to work with and therefore more appropriate for analysing mental wellbeing.

K6 is a more efficient instrument than K-10, but it is not available in SoFIE or the IDI

A 2003 evaluation (Furukawa, RC, Slade, & Andrews, 2003) of the relative efficiency of the K-10 and the K6 concluded that the K-10 may outperform the K6 in screening for severe disorders, but the K6 is preferred in screening for any DSM-IV mood or anxiety disorder because of its brevity and consistency across subsamples.

The K6 is not available in SoFIE, so this was not an option of the current project. However, it may be useful to note the efficiency of the K6 instrument for future research.

Glossary and abbreviations

Table 14 Glossary and abbreviations

Abbreviation	Description
Administrative data	Administrative data is data that government agencies or private organisations collect in conducting their business or services
DHB	District Health Board
IDI	The Integrated Data Infrastructure is a large longitudinal research database constructed by linking administrative and survey data at the individual level. Data is from government agencies, Stats NZ surveys and NGOs
K-10	Kessler Psychological Distress Scale. A checklist aiming to measure depression and anxiety in the last four weeks
MOH	Ministry of Health
MSD	Ministry of Social Development
NGO	Non-Governmental Organisation
NMDS	The National Minimum Dataset is a national collection of publicly and privately funded hospital discharge information, including clinical information, for inpatients and day patients. The NMDS is used by the Ministry of Health, DHBs, PHOs, clinicians, researchers and members of the public for statistical information, clinical benchmarking, and planning and funding
NZHS	New Zealand Health Survey. A survey of the health and wellbeing of New Zealanders used since 1992
PRIMHD	The Project for the Integration of Mental Health Data is the Ministry of Health's single national mental health and addiction information collection of service activity and outcomes data for health consumers. The data is collected from DHBs and NGOs. PRIMHD data is used to report on what services are being provided, who is providing the services, and what outcomes are being achieved for health consumers across New Zealand's mental health sector
SIA	Social Investment Agency
SoFIE	Survey of Family, Income and Employment, collecting information on respondents' work, family, household circumstances, income and net worth
Spine	The IDI consists of a central 'spine' which is the primary person-level dataset (a collection of identifiers) that all other person-level datasets are linked to
Te Kupenga	The Kupenga contains survey data on the social, cultural, and economic wellbeing of Māori in New Zealand, including information about the health of the Māori language and culture.

Te Rau Hinengaro	Te Rau Hinengaro: The New Zealand Mental Health Survey describes the prevalence of mental disorders and their patterns of onset and impact for adults in New Zealand.

References

Andrews, G., & Slade, T. (2001). Interpreting scores on the Kessler Psychological Distress Scale (K10). *Australian and New Zealand Journal of Public Health*, *25*, 494–7.

Carter KN, Shaw C, Collings S, Blakely T, Richardson K. The SoFIE-Health study: Are the results comparable to the New Zealand population? *Australasian Epidemiologist* 2011: 18; 26-32

Carter, P. J., Taylor, B. J., Williams, S. M., & Taylor, R. W. (2011). Longitudinal analysis of sleep in relation to BMI and body fat in children: the FLAME study. *BMJ*, *342*(may26 2), d2712–d2712. https://doi.org/10.1136/bmj.d2712

Deady, M. (2009). A Review of Screening. Assessment and Outcome Measures for Drug and Alcohol Settings. *Network of Alcohol & other Drug Agencies (NADA)*.

Furukawa, T. A., RC, K., Slade, T., & Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine*, *33*(2), 357–62.

Jatrana, S., Crampton, P., Carter, K., & Richardson, K. (2008). SoFIE-Health Primary Care: Overview and Early Results SoFIE-Health Report 3.

Lunt, L. W. (2017). *Preserving the Dignity of the Mentally Unwell: Therapeutic Opportunities for the Criminal Courts of New Zealand*. Retrieved from http://www.fulbright.org.nz/wp-content/uploads/2017/08/LUNT-Preserving-the-Dignity-of-the-Mentally-Unwell-Therapeutic-Opportunities-for-the-Criminal-Courts-of-New-Zealand-.pdf

Ministry of Health. (2016a). Annual Update of Key Results 2015/16: New Zealand Health Survey. New Zealand Health Survey. Wellington: Ministry of Health. Retrieved from http://www.health.govt.nz/system/files/documents/publications/annual-update-key-results-2015-16-nzhs-dec16-v2.pdf

Ministry of Health. (2016b). Population-based Funding Formula Review: 2015 Technical Report. Wellington. Retrieved from http://www.health.govt.nz/publication/population-based-funding-formula-review-2015-technical-report

Morgan, V. A., Waterreus, A., Carr, V., Castle, D., Cohen, M., Harvey, C., … Jablensky, A. (2017). Responding to challenges for people with psychotic illness: Updated evidence from the Survey of High Impact Psychosis. *Australian & New Zealand Journal of Psychiatry*, *51*(2), 124–140. https://doi.org/10.1177/0004867416679738

Oakley Browne, M. A., Wells, J. E., & Scott, K. M. (2006). Te Rau Hinengaro: the New Zealand mental health survey: summary. (M. A. Oakley Browne, J. E. Wells, & K. M. Scott, Eds.). Wellington: Ministry of Health.

Oakley Browne, M. a, Wells, J. E., Scott, K. M., & McGee, M. a. (2010). The Kessler Psychological Distress Scale in Te Rau Hinengaro: the New Zealand Mental Health Survey. *The Australian and New Zealand Journal of Psychiatry*, *44*, 314–322. https://doi.org/10.3109/00048670903279820

Richardson, K., Carter, K., & Hayward, M. (2008). SoFIE-Health Data and Processing Systems SoFIE-Health Report 1.

Social Investment Agency (2018). What you told us. Findings of the 'Your voice, your data, your say' engagement on social wellbeing and the protection and use of data. Retrieved from https://sia.govt.nz/assets/Uploads/what-you-told-us.pdf

Social Investment Unit (2017). Measuring the fiscal impact of social housing services – Technical report, Wellington, New Zealand.

Statistics New Zealand. (n.d.-a). Subnational population estimates (RC, AU), by age and sex, at 30 June 1996, 2001, 2006-16 (2017 boundaries), (subset to 2008 and to those 15 years and over). Retrieved February 16, 2017, from http://nzdotstat.stats.govt.nz/wbos/Index.aspx

Statistics New Zealand. (n.d.-b). Table: Estimated Resident Population by Age and Sex (1991+) (Annual-Jun). Retrieved from http://www.stats.govt.nz/infoshare/

Statistics New Zealand. (2008). Housing problems, Housing Satisfaction and Tenure by Ethnicity, 2008 – GSS. Retrieved from http://nzdotstat.stats.govt.nz/wbos/Index.aspx