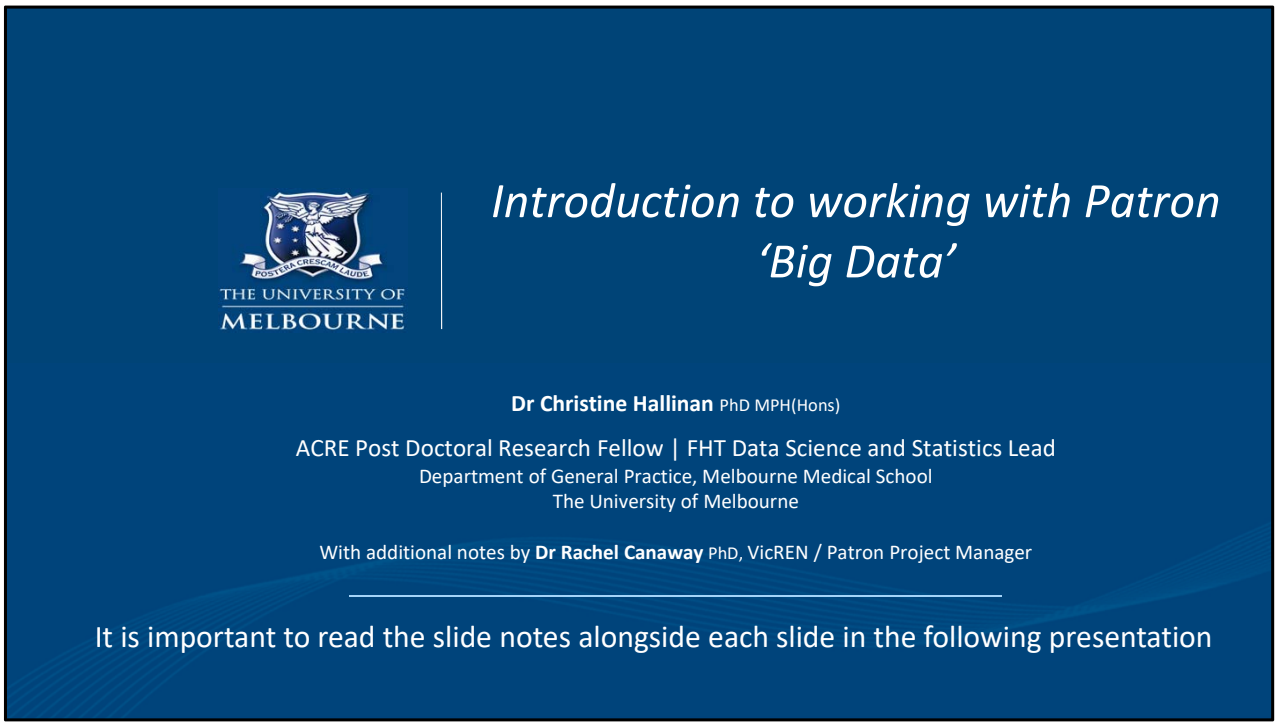


Working with Patron data, a brief presentation provided by Dr Christine Hallinan, Department of General Practice.

Let's start at the critical ending…it's a very good place to start...

AN EXEMPLAR

The use of general practice electronic medical records (EMR) for audit, benchmarking and clinical decision support for patients at risk of CKD or diagnosed with CKD

Patient Gender	Number(n)	Percent (%)
N/A	10	0.11
Male	4,011	43.92
Female	5,111	55.97
Total	9,132	100.00

Current Age (years)	Number(n)	Percent(%)
1-10	1,358	14.87
11-20	779	8.53
21-30	763	8.36
31-40	1,268	13.89
41-50	1,381	15.12
51-60	1,184	12.97
61-70	866	9.48
71-80	837	9.17
81-90	602	6.59
91-102	94	1.03
Total	9,132	100.00

Population	Number(n)	Percent(%)
Not recorded	283	3.10
Aboriginal and Torres Strait Islander peoples	63	0.69
Non - Aboriginal and Torres Strait Islander peoples	8,786	96.21
Total	9,132	100.00

Patients diagnosed with chronic kidney disease	Number(n)	Percent(%)
No	8,624	94.44
Yes	508	5.56
Total	9,132	100

High Systolic BP at risk of chronic kidney disease	Number(n)	Percent(%)
No	7,987	87.46
Yes	1,145	12.54
Total	9,132	100

High BMI at risk of chronic kidney disease	Number(n)	Percent(%)
No	8,895	97.4
Yes	237	2.6
Total	9,132	100

High HbA1c at risk of chronic kidney disease	Number(n)	Percent(%)
No	9,058	99.19
Yes	74	0.81
Total	9,132	100

Patients with CKD prescribed statins and/or ACEARBS	Number(n)	Percent(%)
No	269	53.00
Yes	239	47.00
Total	508	100

- These tables show simple descriptive baseline data that is used to determine which patients in the cohort are at risk of developing CKD and the percent of patients with a coded diagnosis of CKD who have been prescribed a statin.
- The data in these tables was captured using algorithm definitions based on guidelines and recommendations from Kidney Health Australia.
- Patron data does not look like this when you get it. Patron data are comprised of many tables of raw data. Merging of tables is required, but first, data cleaning.
- Patron data stems from general practices that use either Medical Director, Best Practice or Zedmed clinical software systems. At present that data from these different systems are provided in different tables because of the differences in how the data are captured. It is important to have someone (e.g. a GP) familiar with how GP Electronic Medical Record (EMR) data are captured in practice so that data fields can be interpreted correctly. The Patron data book is a work in progress that aims to describe the data fields and their limitations.

When you begin you begin with raw, raw, raw

patient_id	ageat_diagnosis	problem_text	onset_date str23	current_flag
xx00000x000x000x0000x0000x000x0000xxxx00	999	#ed Ribs	2014-12-08 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx01	999	#ered Rib	2007-01-05 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx02	84	renal failure (not otherwise specified)	2016-09-12 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx03	23	#t8;n Bone & Dexa Scan	2006-06-01 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx03	24	'asthma' was It	2007-05-31 17:10:07.000	N
xx00000x000x000x0000x0000x000x0000xxxx04	999	'clotting disorder' - cant recall name .	1800-01-01 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx05	66	'sheep' Story	2007-11-01 15:33:10.000	Y
xx00000x000x000x0000x0000x000x0000xxxx06	45	'white Coat' Hypertension	2006-09-12 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx06	45	(P_ANONYMISED) 's Cyst	2006-10-20 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx07	75	(P_ANONYMISED) 's Cyst - Right	2003-01-01 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx08	999	(P_ANONYMISED) 's Disease	1800-01-01 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx09	999	(P_ANONYMISED) 's Syndrome	2004-02-08 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx10	22	(P_ANONYMISED) As Child	2020-02-13 18:54:18.000	N
xx00000x000x000x0000x0000x000x0000xxxx11	12	(P_ANONYMISED) x2	2007-12-13 18:55:02.000	N
xx00000x000x000x0000x0000x000x0000xxxx12	39	0varian Cyst Resolved	1800-01-01 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx13	999	17mm Gall Stone .	2009-01-10 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx13	999	1st Degree Heart Block	2012-05-07 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx14	44	chronic kidney disease stage 3b	1988-12-14 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx15	81	2 cm liver cyst	2007-01-01 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx16	32	2/05 D+c	2006-04-28 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx17	999	24 hour electrocardi	1800-01-01 00:00:00.000	Y
xx00000x000x000x0000x0000x000x0000xxxx18	69	24h Holter:heart Block/Wenckebach	2005-06-15 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx19	999	2nd Repair Of Paraumbilical Hernia .	2012-03-10 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx20	999	2x BMS in RCA	2015-01-01 00:00:01.000	Y
xx00000x000x000x0000x0000x000x0000xxxx21	35	3 Lap & Dye Tests	2007-12-13 18:54:29.000	N
xx00000x000x000x0000x0000x000x0000xxxx21	40	52 Mm Grade 2 Invasive Lobular Ca R Br	2012-11-02 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx22	66	AAA	2005-01-01 00:00:01.000	Y
xx00000x000x000x0000x0000x000x0000xxxx23	72	AAA - needs yearly ultrasound	2005-01-01 00:00:01.000	Y
xx00000x000x000x0000x0000x000x0000xxxx24	80	Aaa - 3.1 Cm	2007-04-24 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx25	999	Aaa Repair	2009-08-01 00:00:00.000	N
xx00000x000x000x0000x0000x000x0000xxxx26	66	Aneurysm	1997-01-01 00:00:00.000	N

data – uncompressed
obs: 96,388 (rows)
vars:16
size: 40,651,384
variable name
ipracticeno
crs_patient_problem_id
problem_id
patient_uuid
ageat_diagnosis
problem_text
freetext_reasonflag
diagnosis_date
date_first_encountered_conv
date_last_encountered_conv
onset_date
onset_precision
current_flag
problem_criticality
left_right
del_datetime

Now look at this….yes wow, busy slide….that is how I feel when I am working with this data….but there are rewards at the end when you can produce tables like the one that was on the previous slide.

This slide shows a few of the variables that are found the 'Patient Problem' table (read across).

- Patron data is big. Within the entire dataset there are currently over **3 million patients represented**, with **over 1 million of those marked as 'active' or 'current' patients**. A patient is determined as active or current if they have had at least three visits to the clinic within the last 3 years.
- Different practices contributing to the data have different start dates, depending on when they started using the software. Some practices suggest first records start in the 1980s whereas other have first start dates that are much more recent. Patron data recipients will receive only data needed to answer their Ethics and Data Governance committee approved research questions. Use of Patron data outside of the approved scope is not allowed.

Let's look at the data table on the right – there **96,333 observations** that contain data from **16 variable fields** this makes over **1.5 million values**. Look at these variables –some of them purely relate to the functionality of the software (read through a couple)…

Now back to the main table, the rows in the problem text column contains a record of each patient's problem, (ARROW) notice some patients have multiple problems (which makes sense)….

(NO ARROW) What is worth noting here (highlighted in red) is how kidney disease can be recorded in different ways…I have listed all the possible CKD terms in the next slide.
(ARROW) You will see some of the diagnosis ages are recorded as 999 – this reflects missing age data and needs to be changed to a non numerical indicator otherwise the computer may include it in the age stratification table.

Also, look at some of the text in the problem column…it is not always clear

(P_ANONYMISED)….in the case of unclear text sometimes the patient problem becomes more apparent after you merge with other tables such as the pathology table or the prescription table and the clinical table…..

(ARROW) Look at the dates, these need to be transformed so they are readable and so they make programming sense- as cut off dates are required for algorithm development… you will also notice there are some dates that look like the onset was in the year 1800…these dates reflect missing date data, they also need to be converted…

(ARROW) The 'current' flag also is not always correct…(intuitively you would think that a patient with stage 3 CKD in 1988 who has not again visited the practice may not be a current patient)…the same applies to the patient who visited the practice just a couple of weeks ago. Check against prescribing, clinical dates in other tables to determine whether the 'current' flag is accurate… Remember, a patient is determined as active or current if they have had at least three visits to the clinic within the last 3 years – but a new patient can be active without having visited 3 times.

When you sing you continue with more, more, more…

crs_patient_problem_id	freq	crs_patient_problem_id	freq	crs_patient_problem_id	freq
absence of kidney	4	hypertension with nephropathy	1	pyelonephritis	25
acute glomerulonephritis	1	hypertension with renal disease	1	pyelonephritis - left	1
bilateral grade 2 ureteric reflux	1	iga glomerulonephritis	2	pyelonephritis - prob not - left	1
biopsy proven idiopathic membranous	1	iga nephropathy	1	pyelonephritis x2	1
ca kidney, nephrectomy	1	impairment - renal	1	pyelonephritis/septicaemia	1
chronic kidney disease	318	impairment;renal	8	pylonephritis	1
chronic kidney disease - stage 1	3	insufficiency;renal	1	r pyelonephritis	1
chronic kidney disease - stage 2	29	interstitial nephritis	1	reflux nephropathy	3
chronic kidney disease - stage 3	15	kidney disorder	4	reflux ureteric re implant	1
chronic kidney disease - stage 4	9	kidney problem	7	reflux;ureteric	1
chronic kidney disease - stage 5	3	left nephrectomy for renal carcinoma	1	renal biopsy:iga nephropathy/mod - se..	1
chronic kidney disease .	1	left nephrectomy due to congenital	1	renal calculi with pyelonephritis and..	1
chronic kidney disease stage 3a	34	medullary sponge kidney	1	renal disease	4
chronic kidney disease stage 3b	20	nephrectomy	2	renal failure	5
chronic renal failure	63	nephrectomy - cancer	3	renal failure (not otherwise specified)	2
congenital anomaly of the kidney	2	nephrectomy - right	1	renal failure 2008	1
congenital anomaly;kidney	4	nephrectomy staghorn calculus - right	1	renal impairement/microalbuminuria	1
congenital ureteric obstruction	1	nephrectomy;total	2	renal impairment	165
diabetic nephropathy	7	nephritis	3	renal infarct	2
disease;kidney;chronic	1	nephritis - interstitial	1	right nephrectomy	1
disease;renal	2	nephritis admitted	1	right nephrectomy - kidney non functi	1
disorder;kidney	2	nephropathy	4	right nephrectomy following renal haema	1
duplex kidney	7	nephropathy - diabetes mellitus	3	right renal cancer - nephrectomy .	1
duplex kidney - bilateral	1	nephropathy - iga	1	single right kidney with duplex system	1
duplex kidney - left	1	nephrosclerosis	1	slight renal impairment	1
failure;renal;chronic	3	nephrotic syndrome	3	total nephrectomy	8
glomerulonephritis	7	non functioning right kidney	1	unilateral agenesis of kidney	1
glomerulonephritis;acute	1	one functioning kidney - left	1	ureteric reflux	7
horseshoe kidney	5	partial nephrectomy	1	uti recurrent uret reflux-renal failure vesico	3
hypertension renal	1	partial right nephrectomy for benign ..	1	uti ureteric reflux as child	1

Look at all these terms around kidney disease….to get this list from the Patron problem table, you need a basic list of recognised terms so you can tag and pull them from the patient problem field…it is sort of like ‘fishing’ in a really big lake where there are a few big fish you can see….but then there are lots of fish hidden under reeds and in logs that will take time to catch…..

The next few slides just happen to be…

short_descript_prescript	short_descript_prescript	short_descript_prescript
accupril tablets (tablet, coated) 10..	diovan tablets (tablets) 40mg [28]	perindopril arginine 10mg - amlodipin..
accuretic 10/12.5 mg tablets (tablet,..	enalapril actavis tablets (tablet) 1..	perindopril erbumine 2mg oral tablet ..
acetec tablets (tablets) 20mg [30]	entresto 24/26 tablets (24/26 mg) [56]	perindopril generichealth tablets (ta..
adesan hct 32/25 tablets (32/25 mg) [..	exforge 10/160 tablets (10/160 mg) [28]	prexum tablets (tablets) 5mg [30]
apo-enalapril tablets 5mg [30]	felodipine 5mg - ramipril 5mg modifie..	pritor plus 80/12.5 tablets (80/12.5 ..
apo-perindopril arginine/amlodipine 1..	fosinopril sodium 10mg oral tablet 10..	pritor/amlodipine 40/5 mg tablets (40..
apo-perindopril tablets 8mg [30]	gopten capsules (capsule) 4 mg [28]	quinapril 10mg oral tablet 10mg [30]
artane tablets (tablet) 2 mg [200]	idaprex tablets (tablet) 8 mg [30]	ramace capsules (capsule) 10 mg [30]
atacand plus 16/12.5 tablets (16/12.5..	indosyl mono tablets 2mg [30]	ramipril 1.25mg oral capsule (capsule..
avapro hct 150/12.5 tablets (tablet, ..	irbesartan 150mg oral tablet (tablet,..	reaptan 10 mg/5 mg tablets (tablets) ..
avsartan tablets (tablets) 150mg [30]	karvea tablet 150mg [30]	renitec 20 tablets (tablet) 20 mg [30]
candesan combi 32/25 tablets (32/25 m..	karvezide 150/12.5 tablets (150 mg/12..	sevikar 20/5 tablets (tablets) (20/5 ..
candesartan aspen tablets 16mg [30]	lercanidipine hydrochloride 10mg - en..	telmisartan 40mg oral tablet (tablet)..
candesartan cilexetil 16mg - hydrochl..	lisinopril 10mg oral tablet (tablet) ..	teltartan hct 80/12.5 tablets (80/12...
candesartan gh tablets 16mg [30]	micardis plus 40/12.5 mg tablets (40/..	teveten plus 600/12.5 tablet 600mg/12..
candesartan hct gh 16/12.5 tablets (t..	mizart tablets (tablets) 80mg [28]	trandolapril 2mg oral capsule 2mg [28]
candesartan sandoz tablets (tablets) ..	olmertan tablets (tablets) 20mg [30]	triasyn 5.0/5.0 mg modified release t..
capoten tablets (tablet) 25 mg [90]	olmesartan medoxomil 20mg oral tablet..	tritace 2 .5mg tablet 2.5mg (30)
co-diovan 320/25 tablets (320/25 mg) ..	olmetec plus 20/12.5 mg tablets 20 m..	twynsta 40/10 mg tablets [28]
coveram 10 mg/10 mg tablets (perindop..	perindo tablets (tablet) 2 mg [30]	valsartan 160mg - hydrochlorothiazide..
coversyl plus 5/1.25 tablets (5 mg/1....	perindopril an tablets (tablets) 2mg ..	zan-extra 10/10 tablets (tablet, coat..
diovan tablets (tablets) 320mg [28]	diovan tablets (tablets) 40mg [28]	zestril tablets (tablet) 10 mg [30]

Again, the same holds true for the prescription data…however this is more complex because the prescriptions are recorded in free text and often contain spelling errors and incomplete words…this again is like fishing but this time the fish look like old boots (as in not like fish at all)…..

However, what does happen is the more you work with this data…the more familiar you become with the way GPs write, with this familiarity you end up getting a dictionary of prescriptions with common spelling errors and abbreviations which then can be included into the code.

The Patron Management group in the Department of General Practice are convening a **Patron Data Users Group**, a Community of Practice for Patron data recipients to share their experiences in using Patron data. Also, most Recipients of Patron data agree to share data scripts and coding algorithms so that others can build on work that has been done before. The Patron Management group are also working towards a curated data cut that has had appropriate cleaning consistently applied, for use for basic projects, such as student projects. The work involved in creating this is huge and we're not quite there yet.

Oh, no, no....

Consider the algorithm for a coded diagnosis of hypertension

At least 2 most recent consecutive Systolic blood pressure (SBP) > 140 OR at least 2 most recent consecutive Diastolic blood pressure (DBP) > 90

...this is NOT consecutive readings it is consecutive visits

patient_id	measure_id	ageat_measure	measure_time	glookup_crs_measure_type_measure	measure_type_id	measure_value
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2205483	67	4-May-17	SYSTOLIC	12	145
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2203484	67	4-May-17	DIASTOLIC	13	67
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2206966	67	5-Jul-17	SYSSTAND	15	165
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2206967	67	5-Jul-17	SYSTOLIC	12	135
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2206968	67	5-Jul-17	DIASTOLIC	13	97
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2206969	67	5-Jul-17	SYSSTAND	25	145
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X00	2206970	67	5-Jul-17	DIASSTAND	26	89

patient_id	measure_id	ageat_measure	measure_time	glookup_crs_measure_type_measure	measure_type_id	measure_value
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X01	2206554	72	17-Jul-17	SYSTOLIC	12	154
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X01	2206555	72	17-Jul-17	DIASTOLIC	13	80
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X01	2206556	72	17-Jul-17	SYSTOLIC	12	142
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X01	2206557	72	17-Jul-17	DIASTOLIC	13	73
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X01	2206558	72	17-Jul-17	PULSE	11	74

Look at the tables in this slide…to pull data within the constraints of the hypertension algorithm you would expect to write a code which 'tags' the last two systolic BPs that were above 140…however in the case of the first table…(ARROW) the last three systolic BPs were taken on the same day…hence the last two high Systolic BPs would be pulled from the first and the fourth most 'recent' BP reading (not the first and the second)….

(ARROW) In the second table issue of repeated measurements on same day…

and, there's more….

Consider the algorithm for a diagnosis of CKD = two abnormal readings of Albumin Creatinine Ratio (ACR) recorded more than 3 months apart
Note: test for an abnormal Urine ACR and/or eGFR refers ONLY to the most recent three (3) test results, up to a maximum of 3 years from the current date.

Consider the algorithm for a coded diagnosis of diabetes = at least 2 consecutive HbA1c >= 6.5% (Note: difference in values between HbA1c and HbA1c_IFcc).
Note: test for high HbA1c refers ONLY to the two most recent consecutive tests.

X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	measure_id	ageat_measure	measure_time	glookup_crs_measure_type_measure	measure_type_id	measure_value
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128391	69	08-Mar-2019	TRIG	19	1.8
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128392	69	08-Mar-2019	HDL	20	1.5
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128393	69	08-Mar-2019	LDL	21	3.3
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128394	69	08-Mar-2019	eGFR	11007	13
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128395	69	08-Mar-2019	CREATININE	22	76
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128417	69	08-Mar-2019	UCREATININE	11018	12.5
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128417	69	04-June-2019	ACR	11018	44.9
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128419	69	08-Mar-2019	ACR	11002	39.2
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128420	69	08-Mar-2019	HAEMOGLOBIN	11011	142
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128421	69	08-Mar-2019	HBA1C_IFCC	11012	75
X000X0X0X-00XX-XXXX-X0X0-XX00XXX00X02	1128422	69	08-Mar-2019	HBA1C	11011	12.7

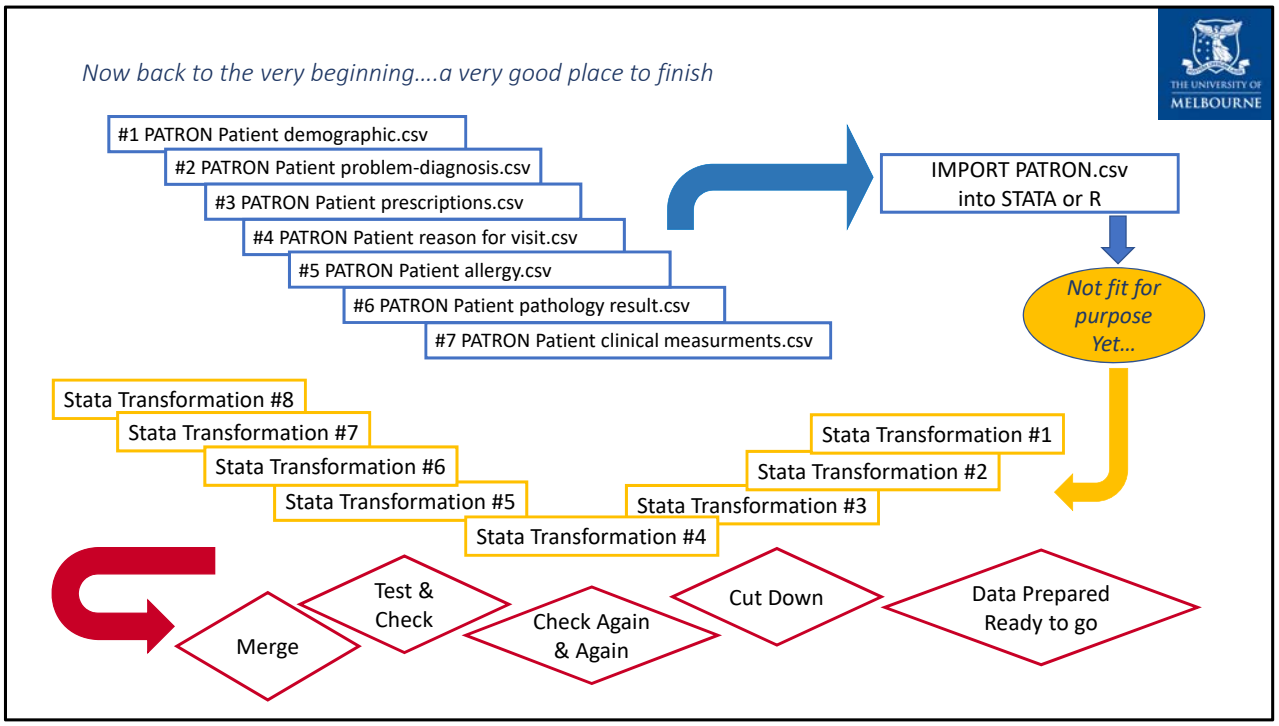
Again, for the algorithms in this slide you need to write a code that 'tags' the last two abnormal ACRs and HbA1c's that have a reading >=6.5……

This table highlights a few issues…
(i) (ARROW) two HbA1C variables (HBA1C_IFCC [20-42 mmol/mol] **International Federation of Clinical Chemistry (Europe)** & HBA1C [4%-5.6%) – old percentage value

(ii) (ARROW) multiple creatinine variables - **creatinine** [**serum creatinine**≥ 25 µmol/l within 48 hours]; **ucreatinine** [urine creatinine]; and ACR–**urine albumin creatinine ratio** (uACR) - **urine ACR >25 mg/mmol (males) or >35 mg/mmol**
(females) provide measurements of the same indicator (HbA1C & creatinine) -but use different units , again we have the issue around multiple readings on the same day

(iii) also this slide illustrates constraints relating to the time factor in algorithm development [i.e. the requirement for diagnosis of renal impairment to be two abnormal readings of Albumin Creatinine Ratio (ACR) recorded more than 3 months apart]

(iv) …here we can see there is evidence of two creatinine values above the recommended range, however they will not be tagged as abnormal because they
were recorded within (under) the three month threshold (2 months and 27 days).



- To begin, import that different Patron data .csv tables into the data analytics package that will be used – e.g. STATA or R.
- Merge tables.
- Test and check – cleaning as you go
- Check again and again
- Cut down data so it can be prepared and ready to analyse

Let's end at the very end….PATRON data = transformation & modelling = output = simple descriptive statistics…

Patient Gender	Number(n)	Percent(%)
N/A	10	0.11
Male	4,011	43.92
Female	5,111	55.97
Total	9,132	100.00

Current Age (years)	Number(n)	Percent(%)
1-10	1,358	14.87
11-20	779	8.53
21-30	763	8.36
31-40	1,268	13.89
41-50	1,381	15.12
51-60	1,184	12.97
61-70	866	9.48
71-80	837	9.17
81-90	602	6.59
91-102	94	1.03
Total	9,132	100.00

Population	Number(n)	Percent(%)
Not recorded	283	3.10
Aboriginal and Torres Strait Islander peoples	63	0.69
Non - Aboriginal and Torres Strait Islander peoples	8,786	96.21
Total	9,132	100.00

Patients diagnosed with chronic kidney disease	Number(n)	Percent(%)
No	8,624	94.44
Yes	508	5.56
Total	9,132	100

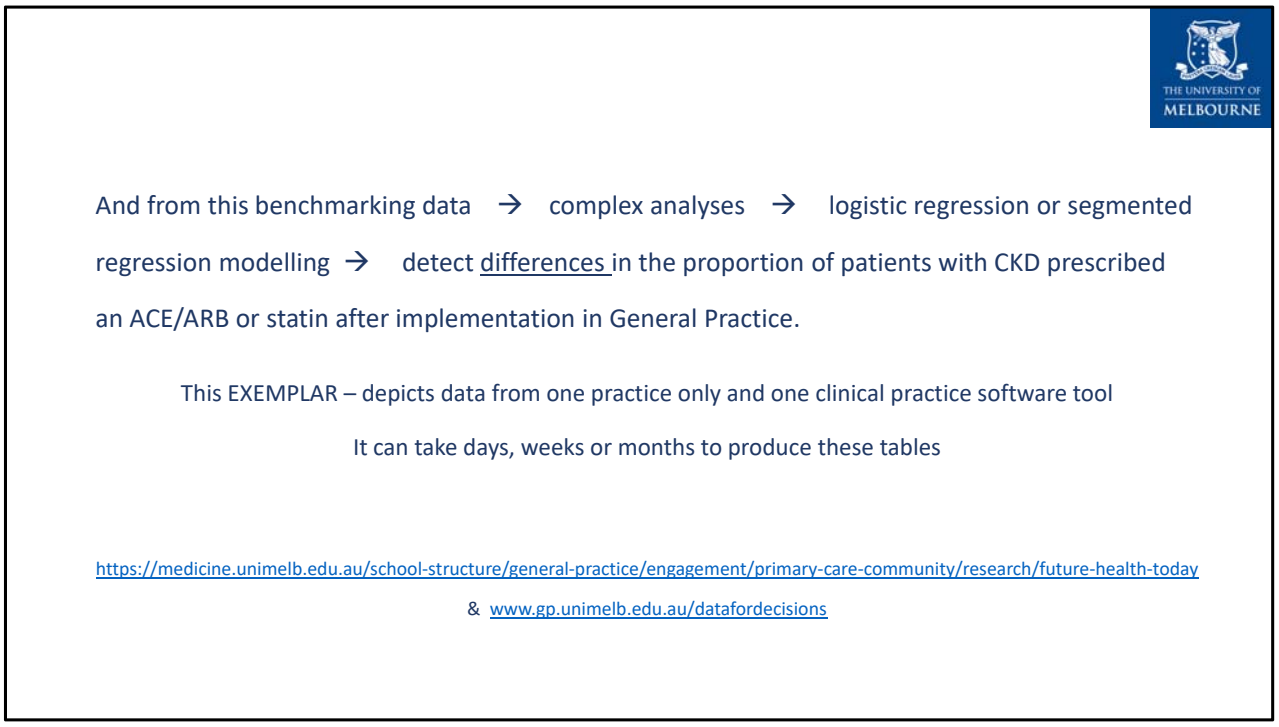
High Systolic BP at risk of chronic kidney disease	Number(n)	Percent(%)
No	7,987	87.46
Yes	1,145	12.54
Total	9,132	100

High BMI at risk of chronic kidney disease	Number(n)	Percent(%)
No	8,895	97.4
Yes	237	2.6
Total	9,132	100

High HbA1c at risk of chronic kidney disease	Number(n)	Percent(%)
No	9,058	99.19
Yes	74	0.81
Total	9,132	100

Patients with CKD prescribed statins and/or ACEARBS	Number(n)	Percent(%)
No	269	53.00
Yes	239	47.00
Total	508	100

Finally, the descriptive data tables emerge.



'And from this data more complex analyses can be undertaken such as logistic regression or time series analysis (straight arrow) so that we can detect differences in clinical outcomes …. for example the proportion of patients with CKD prescribed an ACE/ARB or statin …. (curved arrow) which can be then fed back to the general practices….'

How long would it take to get these benchmark tables?
'What comes with this work is an increased understanding of the characteristics of the data. It's an iterative process so as more data comes in more coding is developed so will working with data will be quicker and more efficient'

Coded diagnosis of CKD- in the presence of hypertension, albuminuria or high absolute cardiovascular risk
ACE inhibitor and/or an angiotensin 2 receptor antagonist and statin therapy,….which is viewed as best practice by Kidney Health Australia

Medical knowledge of disease and prescriptions, and knowledge of how data are captured in general practice, plus plenty of patience and willingness to be thorough are needed if results from Patron data are to accurately reflect what is happening with patients in primary care.

Thank you

Dr Christine Hallinan PhD MPH(Hons)

DDQI Data Science and Statistics Lead
Department of General Practice
Melbourne Medical School
Faculty of Medicine, Dentistry & Health Sciences
The University of Melbourne