Validating the Portal Population of the UK MS Register

Authors: Middleton RM, Rodgers WJ, Chataway J, Schmierer K, Akbari A, Tuite-Dalton K, Lockhart-Jones H, Griffiths D, Noble DJ, Jones KH, Al-Din A, Craner M, Evangelou N, Galea I, Harman P, Harrower T, Hobart J, Husseyin H, Kasti M , Kipps C, McDonnell G , Owen C, Pearson O , Rashid W, Rog D, Wilson H, Ford DV

The UK Multiple Sclerosis Register (UKMSR) is a large cohort study designed to capture 'real world' information about living with multiple sclerosis (MS) in the UK from diverse sources. The primary source of data is directly from people with Multiple Sclerosis (pwMS) captured by longitudinal questionnaires via an internet 'portal'. Clinical Data is also captured from MS Specialist Treatment centres across the UK, a proportion of the internet population has their MS diagnosis verified by their Neurology team at these hospitals.

Key questions of the MS Register are how representative of the UK MS population is the internet component and how are their diagnosis of MS validated.

We analysed the MS Register for a number of key characteristics. The internet (n=11,021) and clinical (n=3,003) populations were studied for key shared epidemiology. We found them to be closely matched for mean age at diagnosis (clinical=37.39, portal=39.28) and gender ratio (female %, portal=73.1, clinical=75.2). Using the non-parametric two sample Kolmogorov-Smirnov test in order to discover if the different continuous variables examined are from the same distribution resulted in the null hypothesis being rejected only for age at diagnosis ($D = 0.078$, $p << 0.01$). With the populations therefore being drawn from different distributions. In all other analysis performed the populations were shown to be drawn from the same distributions.

Our analysis has shown that the MS Register portal population is highly analogous to our entirely clinical (validated) population and therefore for research can be utilised as a viable and valid cohort to study.

Introduction

Multiple sclerosis (MS) is an inflammatory demyelinating and degenerative disease of the central nervous system (CNS), and the most common non-traumatic cause of disability in young adults worldwide (1). The dominant phenotype is characterised by relapses (attacks) and remissions, relapsing MS (RMS). This evolves at 10-15 years into secondary progressive MS (SPMS) in the majority of those affected. About 15% of people with MS (PwMS) develop progressive neurological dysfunction from onset – primary progressive MS (PPMS).(3)

The disease affects about 120,000 people in the United Kingdom (UK) with an incidence that appears to be increasing by approximately 2.5% per annum (4). There is a significant societal burden of £3bn/year (5)

In order to comprehensively map the prevalence and characteristics of MS across the UK, the MS Society of Great Britain & Northern Ireland commissioned Swansea University Medical School, home to the SAIL Databank (6), to develop the UK Multiple Sclerosis Register (UKMSR) in 2010. During a pilot phase (36 months) data from various sources (4 NHS Trusts and from people with MS) were acquired to carry out data linkage and develop a 'real world' dataset. From 2011, we started to capture original data drawing on the experience from a number of European Registries(7) to develop a joint dataset consisting of data reported by (i) participants and (ii) health care professionals (HCP). Hence, altogether three data sources were combined:

1) Data provided by PwMS via the UKMSR's internet portal ('portal data').
2) Data reported by HCP either using a hardcopy datasheet, or a summary spreadsheet ('clinical data').
3) Data mined from general practice and inpatient hospital records ('routine data').

A key aim of the UKMSR is for these data sources to be linked longitudinally, thereby providing a comprehensive cohort that can be pseudonymised and made available for academic research and audit, subject to appropriate governance.

The UKMSR collects a standardised minimum clinical dataset [Appendix 2] which was developed by its Clinical Advisory Group (CAG), comprising of UK neurologists with an interest in MS. The clinical dataset is a pragmatic annual collection, following informed consent, of demographic as well as basic clinical indices. This clinical data can then be linked with portal data. Linkage occurs either when the participant enters their Study ID, or through deterministic and probabilistic methodologies(8).

Other studies using solely participant supplied data such as the North American Research Committee on Multiple Sclerosis (NARCOMS) were successful in validating their populations

by carrying out expert review on a sample of their population. Their observed eligible population for this validation exercise was 142 participants out of 30,691 (9) following careful removal of their inactive population, and participants who did not respond. Despite some caveats, NARCOMS reported the diagnostic accuracy of their population as being 98.7±1.3% in 2006. For the MS Register every patient that is consented at an NHS site, by design has a McDonald criteria confirmed diagnosis of MS by a UK Neurologist. Therefore, there is no need to validate this element of the population.

Methods

Study design and participants

The UKMSR consists of (i) a repository of patient-reported demographics and outcome measures (PROMs) for self-declared pwMS, and (ii) a dataset based on patient data collected by clinical centres. PwMS giving informed consent for collection and reporting of their clinical data are encouraged to also sign up to the portal. As a result, validation of PROMs against clinical data provided by a UK neurologist becomes feasible.

Figure 1 illustrates the two types of data stored by the UKMSR, and the overlap between the two thereby enabling validation of PROMS against clinical data provided by UK neurologists.

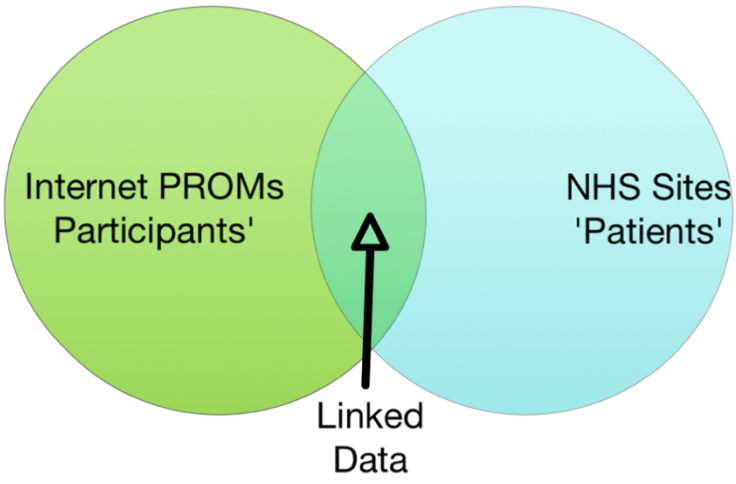

Figure 1- Venn diagram showing the linked population as an intersection of the internet participants with consented patients from a clinical site.

Portal data

pwMS become aware of the UKMSR through a variety of sources including word of mouth – from their friends or colleagues, information at clinical sites, advertising by the MS Society, social media, blogs or attendance of a presentation by UKMSR staff. Participants are over 18 years of age and resident within the United Kingdom. They agree to Terms of Service (10) stating the Register's responsibility for the use and storage of their data. They are asked to enter demographic and descriptive data about their diagnosis, medical and family history as part of a 'baseline questionnaire'. Users are then invited to complete a number of PROM questionnaires. Participants are subsequently reminded every three months via email to return to the portal and complete these questionnaires, thereby building up a patient reported narrative that evolves over time. The number of PROMs assessed has increased and now includes the Hospital Anxiety and Depression Scale (HADS) (11), EurQol Quality of Life (EQ-5D-5 or 3L?) (12) and the Multiple Sclerosis Impact Scale (MSIS29v2) (13) Up to five years of longitudinal data in some cases are now available.

Clinical data

pwMS over the age of 18 whose diagnosis has been confirmed by a neurologist (14) are invited to give informed consent for sharing their clinical data with the UKMSR. Consent can be given face to face, by post or by eConsent application. The UKMSR has been approved as a study by the South West Central Bristol Research Ethics Council, initially under registration code 11/SW/0160. Approval was renewed after five years under registration code 16/SW/0194. Informed consent is usually followed, in the same session, by collecting a 'minimum data set'. [appendix 2]. This dataset comprises demography, disease history and course, onset symptoms, current symptoms, relapse and progression information as appropriate, current and previous disease modifying therapies (if appropriate) and 2 measures of function, a timed walk and clinician supplied EDSS score.

- 8 Sites use the iMed MS Clinical System
- 2 Sites utilise another dedicated MS Clinical System
- 14 return an Excel spreadsheet based Case Return Form (CRF)

Participating clinical centres [appendix 1]consent patients on an ongoing basis, and this data is transmitted to the UKMSR using encrypted secure transfer methods. In order to facilitate recruitment data capture is encouraged using any secure system, including electronic health records and paper recording. Data transfer to the UKSMR from the clinical centres usually takes place once/month.

Data

The data contained in UK MS Register is stored across a number of different Microsoft SQL Server 2014 Databases. The data collected for this study were initially filtered and linked from

these different databases using Structured Query Language (SQL) queries. Initial analysis, data cleansing and aggregation was also conducted using SQL in the MS SQL Server environment.

Statistical Analysis

Analysis was conducted using the R statistical computing programming language (15) in the RStudio environment. A comparison between the populations in the portal and clinical data sets was done in the first instance using simple descriptive statistics: mean (standard deviation) for continuous data and frequency tables for categorical data on key demographics markers. A two-sample Kolmogorov-Smirnov test was then implemented using age at diagnosis and current age. This is a non-parametric statistical test which determines if two different continuous variables are from the same distribution.

Portal Data

The initial interrogation of the portal data in March 2017 indicated that there were 14,720 'unique' registered participants with a valid email, date of birth and gender. Of these, 77 had elected to leave the study by clicking the 'leave study' link, and a further 1,648 who had either died (informed by family member/carer) or 'merged' their records due to opening multiple accounts under different email addresses. Leaving 13,072 participants with the initial baseline criteria at registration. Furthermore, only the most recent responses were used (as updates can be made). Those participants who had not indicated year of diagnosis and type of MS at diagnosis were also removed from the result set leaving 11,051 eligible individual records. This number was further reduced by removing those records with obvious input errors such as diagnosis year in 1927, or date of birth and year of diagnosis having matching dates; leaving 11,021 entries.

Clinical System

Clinical data is submitted via clinical system (iMed) or by CRF (Microsoft Excel spreadsheet). 8 iMed sites have submitted 2,306 patients to the MS Register, of these 163 had missing or invalid data, leaving 2,143 valid records to examine. For the Excel sites 888 records for 7 sites have been submitted and when tested against the same criteria as for iMed 860 remain, leaving a combined total of 3,003. Figure 2 shows the consort diagram for the overall process.

Consort Diagram

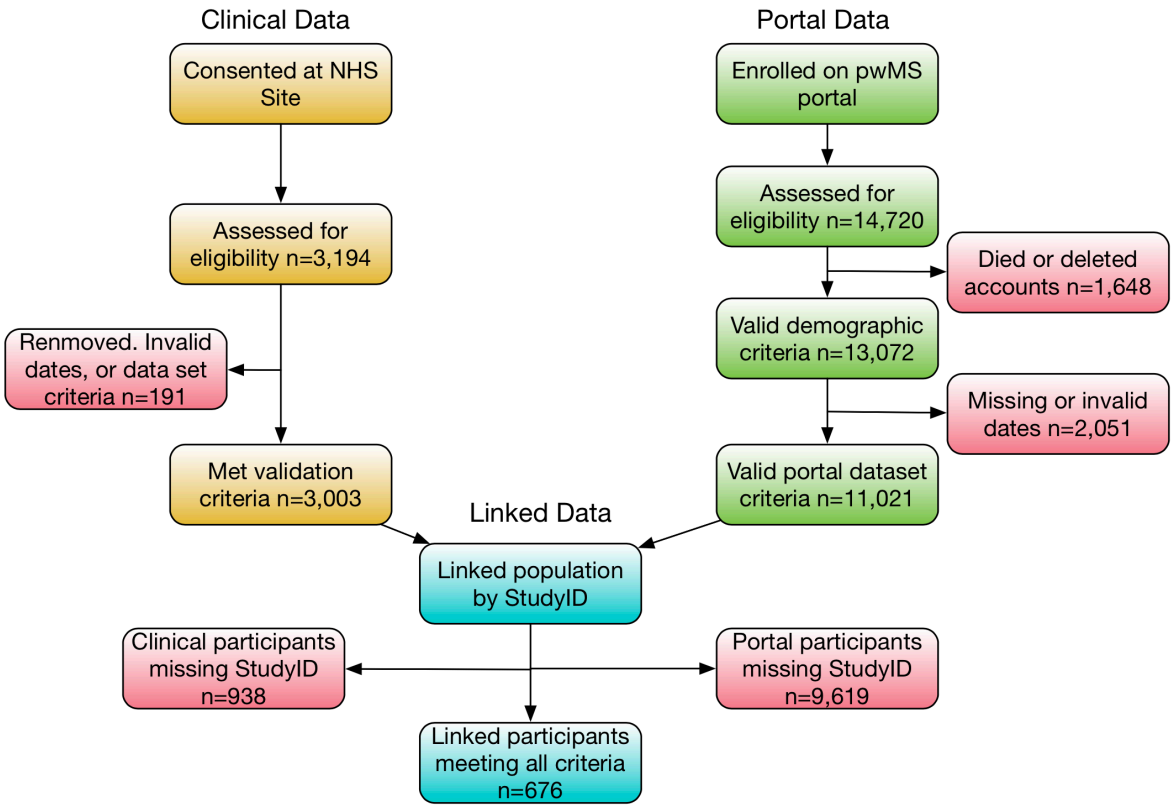

Figure 2: Consort diagram showing data selection criteria

Linked

The linked datasets look at those data that intersect between the clinical and portal populations, through confirmed presence of a studyID. For the Excel sites, of the 860, 9 were missing the code – Leaving 851. For the iMed sites, 2,705 records were checked using the same criteria leaving 1,776 (929 removed).

For the portal participants - using the same criteria as above, the 11,021 entries checked for the study ID on their records. This reduced the number who 'could' be linked from the internet population to 1,402 (9,619 removed).

Once positive linkage was made (allowing for nulls and incorrect values in diagnosis date on the portal side) there were 676 participants in the clinical and portal data.

Results

The MS Register published a paper after one year of operation in 2012, describing the portal population (16), where 7,279 people had registered. Now this number has almost doubled to 14,720.

Of these, 13,072 people have provided their location, gender and MS Type. And 11,051 have additionally provided a date of diagnosis. The location of MS Register participants is highly analogous to the overall UK population as can be seen in Table 1.

Country	MS Register Portal % population	MS Register Clinical population	Entire UK population
England	77.60	75.63	84.14
Northern Ireland	4.26	12.36	2.84
Wales	7.71	12.01	4.76
Scotland	10.43	0	8.25

Table 1 : MS Register population distribution compared to the UK general population

The clinical population is slightly less geographically diverse – The MS Register currently has no clinical sites within Scotland. Scotland has a separate MS Incidence Register (17) that all Scottish NHS Trusts are required to participate with.

6,092 participants have given informed consent to be part of the UK MS Register as of March 2017, as with all 'real world' data there is a 'lag' in data quality that the Register reports on and returns to the originating site to manage. That is, data that is entered must be validated and returned to the originating site to have data entry errors corrected. There is a population of people on the portal who have never provided more than an email address. At launch this figure stood at 1,904. By March 2017 this has risen to 2,686.

Characteristics of the Register populations

Age

Table 2 shows the results of the three datasets by age, age at diagnosis and MS type at diagnosis.

	Clinical	Portal	Linked
	n=3,003	n=11,021	n=676
Age (mean)	48.8 ± 11.9	52.3 ± 11.7	48.3 ±11.3
Age at diagnosis	37.4 ± 10.6	39.3 ± 10.2	38.6 ± 10.6
Gender (female)	2,178 (75.2%)	8,052 (73.1%)	493 (72.9%)
PPMS	198 (6.5%)	1,514 (13.7%)	51(7.5%)
RRMS	2,521 (83.9)	7,408 (67.2%)	567(83.8%)
SPMS	122 (4.0%)	839 (7.6%)	21(3.1%)
Other	119(5.3%)	1,260 (11.%)	37(5.4%)

Table 2: MS Register datasets compared by age, age at diagnosis and MS Type at diagnosis

The similarities between the three captured populations are notable, taking the clinical population as the 'gold standard' it establishes the gender, age and age at diagnosis ranges for PwMS along with the MS Type at Diagnosis.

Comparing the ages of both the portal and clinical data using the two-sample K-S test revealed the populations were for the most part drawn from different distributions. The null hypothesis was rejected in the comparison of the datasets for both the current ages (D = 0.131, $p << 0.01$) and the ages at diagnosis (D = 0.078, $p << 0.01$). This means that the populations are drawn from different distributions. As the D statistic is small, the overall difference in the age distributions is minimal [Graph x and Graph z]. When stratifying the age populations by disease type equally RRMS the null hypothesis is still rejected, but by a much smaller margin (D = 0.131, $p << 0.01$). Figure 3 shows the overall kernel density of the MS Register clinical and portal populations for age at diagnosis, with Figure 4 highlighting the same data for people with relapsing remitting disease.

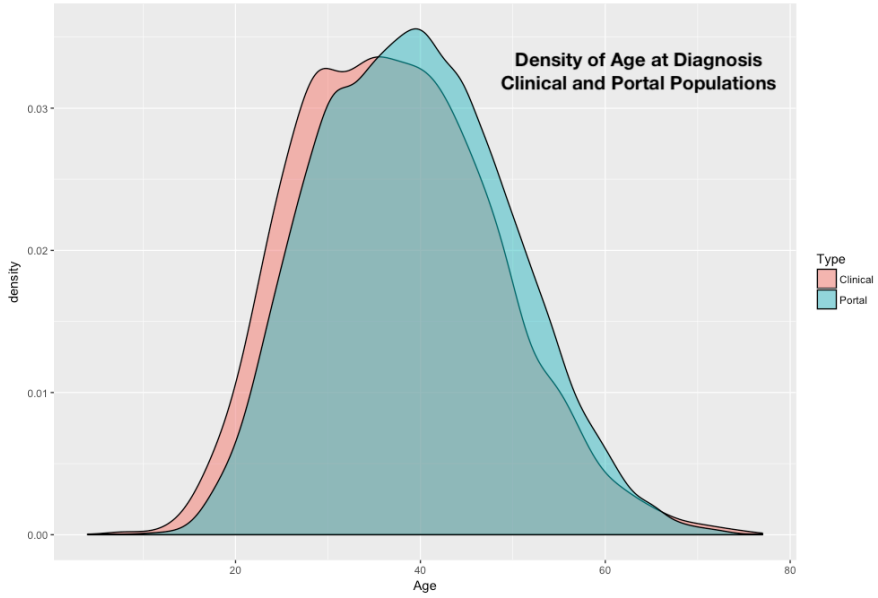

Figure 3 Kernel density of age at diagnosis, Portal and Clinical Population

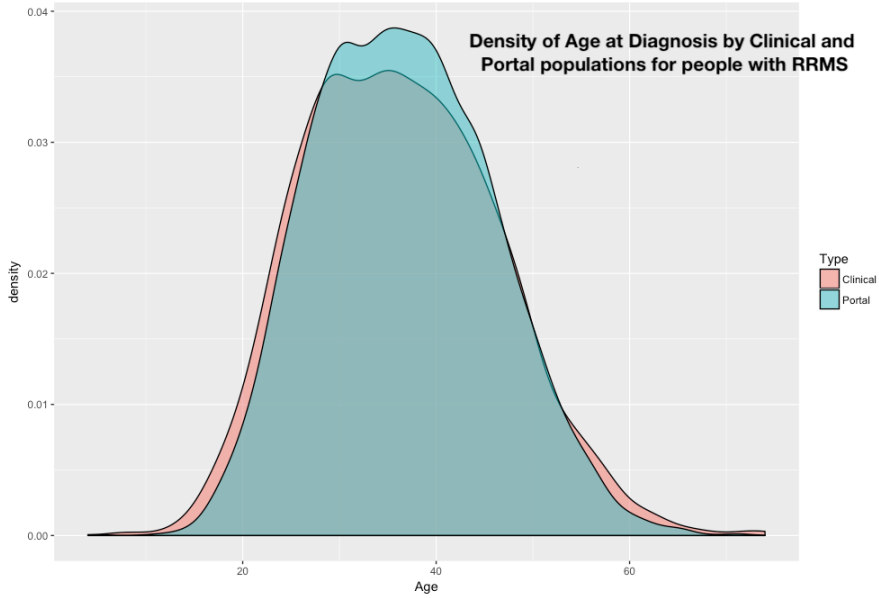

Figure 3 Kernel density of current age at diagnosis, Portal and Clinical Population for people with RRMS

Gender and MS Type

There is little in the literature about exactly how a 'general' population of MS is stratified by disease type Clinical Registers tend to recruit for specific clinical trials or only recruit those patients that attend at clinic. There is a potential loss to follow up for those pwMS who don't attend a clinic. The LORSEP registry of MS indicated that 87% of its 2,871 patients were diagnosed with RRMS at onset (18) and this is borne out in similar cohorts worldwide (19) (20). The MSIF Atlas of MS (21) defines RRMS incidence at >81% for most of Western Europe. Both the clinical and linked datasets above show RRMS at diagnosis as >70% with the portal population at 67% being only slightly under this threshold. There are small differences between the populations declaring SPMS, Other and PPMS as observed on the portal.

In 2012 the Portal element of the MS Register reported that 15% of the population had Primary Progressive MS (PPMS), 63% Relapsing Remitting MS (RRMS) and 8% having Secondary Progressive MS (SPMS). 14% stated that they did not know what sort of MS that they had been diagnosed with(16) . Although overall numbers have increased since then the trend is towards harmonising with the clinical population.

The clinical group is on average 10 years older than those involved in the portal- though still within the standard deviation of the groups. However, MS is a disease that is classically diagnosed in the mid to late 30's (22) and the date of diagnosis for both domains are within 0.07 of each other.

It has been observed in the literature, that the gender distribution in people with MS has become increasingly more female over the last two decades (23) (24) (25) Register data from the clinical and portal elements as seen in Table 3 support this.

	Total Linked	Female %	Total Portal	Female%	Total Clinical	Female %
PPMS	50	52%	1514	52.4%	198	48.4%
RRMS	567	74.8%	7408	77.7%	2,521	74.6%
SPMS	21	76.2%	839	66.3%	122	66.3%
Other	37	72.9%	1260	74.4%	162	73.4%

Table 3: Gender distribution across the UK MS Register

The data across all areas of the Register remain consistent with what would be expected from a cohort of people with MS, this is especially interesting in the male population with PPMS, where clinically more males are present that females(26) and the percentage of females remains low across the other disease types. This is in line with other large clinical database led studies (27) .

Quality of Reporting

It is here that the differences in the linked dataset begin to become more distinct, with the 'linked' population being a stepping stone between the self-reported and clinically confirmed. As can be seen in Table 4 there were strong matches in gender and date of birth – These values should be given and can almost certainly classified as transcription errors. Just over 80% of clinicians and PwMS agree on the type of MS at diagnosis: this is potentially down to participants selecting the type of MS that they have now, and not at time of diagnosis. The overall proportion of confirmation between this remains high. This remains correspondingly strong in memory of dates of diagnosis, with 62% of participants remembering the year of diagnosis exactly and this percentage rises to 81.7% when a 1-year margin of error is allowed. For a population that have a mean of 13 years since diagnosis, this could be seen to be an acceptable margin of error.

Fields	Count	%
Total Records	667	100
Matched Gender	664	99.55
Matched DOB	652	97.7
Matched Year of Diagnosis	430	64.47
1 Year within diagnosis	116	17.39
2 Years within diagnosis	37	5.55
3 Years within diagnosis	20	3.00
>3 Years from Diagnosis	37	5.55
Errors in Diagnosis Date	27	4.05
Matched Type of MS	549	82.31

Table 4: Quality of self-reported variables against clinically reported data.

Discussion

There are clearly many potential issues with a general observational Multiple Sclerosis Register when compared to more focused clinical trial databases or even dedicated clinical management systems. In these modes, it is easy to capture 100% of a, small identified population and monitor specific changes. Itit is highly unlikely that a 1:1 mapping of clinical consent to participant sign up via the Register will ever be achieved – due to the MS Register not being a requirement at every NHS site across the UK and there being patients who no longer attend a designated hospital for MS treatment of symptom management. Therefore, the validation of the online cohort is important.

The UKMS Register design was from the outset to be pragmatic and work with systems that were already in place within the NHS, and capture a dataset from the online population that would have the 'most' utility to clinicians, academics and people with MS. To that end many of the criteria that a dedicated 'study' would have to analyse are sparsely populated or just not feasible to capture in a large general-purpose instrument.

Many large IT projects involving the NHS have failed due to trying to implement large scale cultural change, of note is the failure of the National Programme for IT (28). Subsequent health care delivery within the NHS has been to build on existing platforms, and mindful of these issues, the Register has taken this pragmatic approach.

A core reason for establishment of the Register was to look at the types of MS that people are diagnosed with, as there was a lack of UK data The reasons for this are well established with neither inpatient hospitals systems capturing more than the IDC10 code (G35-G37)

"Demyelinating disease of the Central Nervous System" (29) and historically general practitioner systems via READ code as "Multiple Sclerosis"(F20) (30) – We feel that given the good quality data supplied from the portal population and as validated by the clinical population, that a more accurate picture of the disease types, duration since diagnosis and average age across the UK emerges

What is striking about these data, is how well it maps across the 3 domains discussed here and with the evidence in the published literature. If the goal of the MS Register was to have an online population that is statistically similar to a clinical MS service, then it has generally succeeded.

One limitation is the lack of clinical data from Scotland as described above and it would be important to compare the prevalent population of the UK MS Register with Scottish Register

When looking at the ages of participants, there are some differences between the clinical and portal population. The latter is slightly younger, but this is true of Internet usage as a whole in the UK (31) Some research has suggested that Internet use is decreased amongst people with chronic disease (32) but in the UK Internet use amongst all age groups is increasing (33) with only 10.2% of the population never having accessed the Internet. The MS Register only captures age at confirmed diagnosis, rather than age at onset, though this definition varied in the literature and other studies (34), (35) Fundamentally when comparing age at diagnosis of the Internet population , it aligns well with other studies. (4), (20). There is only a 4-year difference between the mean age of the portal population against that of the clinical and linked. What is more interesting is that the mean age at diagnosis across the populations is within 1 year of each other. This is more statistically significant than the current age of the individuals.

The recruitment of 'self-declared' people with MS via the internet is an inherent limitation, there will be a degree of selection bias, its likely we only get the proportion of the population that is interested in taking part in research. This is somewhat balanced by having 'general' recruitment at a variety of neurology centres across the UK where all patients can be encouraged to sign up. In looking at the types of MS that people are recruited with from clinics however there seems to be a higher than expected degree of people with RRMS – Leading to a suggestion that the majority of patients may be recruited at Disease Modifying Therapy clinics. To this end the Register is attempting to recruit more community care NHS Trusts and to encourage exiting sites to make more use of the postal consent methodology to capture those patients that may no longer attend clinics.

Conclusions and future work

The results above show that in comparison to a 'clinical' population it does appear that there is representative sample beginning to emerge in the online 'portal' population. This paper is a firm first step in being able to 'treat' that online cohort of people with MS as a valid one and enhances the research that has already been carried out via the MS Register.

For the future the MS Register will continue to improve the quality and quantity of data collected from clinical systems and the internet by enhancing the validation of data entry fields and checking the logic of entered data before storage. We will also seek increased validation by linkage with other routine datasets, where such data is available. We have already improved our capture of data from NHS sites by transitioning to a new eCRF system that is easier to access from within the NHS and provided instant feedback on collected data. We will carry out similar improvements on the Internet collection methodology and look into provision of data in new ways – Such as smartphone 'apps'. Further validation techniques will be investigated, and we will carry out a study where researchers will access patient notes in an established (pilot) Register site and then a carry out a comparison between notes and the internet data from a more recently added NHS Site

There does appear to be a 30% uptake of patients recruited at clinical sites that then go on to use the Internet – This number varies between sites slightly but is relatively predictable. We therefore need to recruit more patients clinically and attempt to make use of novel linking and data capture techniques when possible. We will also look to make data capture more seamless by applying novel techniques, such as the application of Natural Language Processing software. We would also like to look at carrying out similar studies with other UK Data collections such as the Scottish MS Register and the SWIMS study.

The methodology of data capture from the internet, clinical systems and routine data has proven to be useful and crucially it is patient centred. Having established the validity of the cohort it becomes possible to make use of this research in a variety of ways. The first will be selecting subsets of the cohort that may be appropriate for clinical trials. Of additional interest will be testing the whole cohort with novel online outcome measures that would be difficult to test on less well-characterised or smaller clinical cohorts. For example, how a web-based participant supplied EDSS score compares to a formal clinically supplied one.

Acknowledgements

The UK MS Register is entirely funded by the MS Society. Award 64

References

1. W. J. Brownlee, T. A. Hardy, F. Fazekas, and D. H. Miller, "Diagnosis of multiple sclerosis: progress and challenges," *The Lancet,* vol. 389, pp. 1336-1346.

2. De Stefano N, Giorgio A, Battaglini M, Rovaris M, Sormani MP, Barkhof F, et al. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. Neurology. 2010 Jun 8;74(23):1868–76.

3. Leray E, Yaouanq J, Le Page E, Coustans M, Laplaud D, Oger J, et al. Evidence for a two-stage disability progression in multiple sclerosis. Brain. 2010 Jul 1;133(7):1900–13.

4. Mackenzie IS, Morant SV, Bloomfield GA, MacDonald TM, O'Riordan J. Incidence and prevalence of multiple sclerosis in the UK 1990-2010: a descriptive study in the General Practice Research Database. J Neurol Neurosurg Psychiatry. 2014 Jan 1;85(1):76–84.

5. Thompson AJ. Challenge of progressive multiple sclerosis therapy: Curr Opin Neurol. 2017 Jun;30(3):237–40.

6. Ford DV, Jones KH, Verplancke J-P, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Health Serv Res. 2009;9(1):157.

7. Flachenecker P, Stuke K. National MS registries. J Neurol. 2008 Dec;255(S6):102–8.

8. Sayers A, Ben-Shlomo Y, Blom AW, Steele F. Probabilistic record linkage. Int J Epidemiol. 2016 Jun;45(3):954–64.

9. Marrie RA, Cutter G, Tyry T, Campagnolo D, Vollmer T. Validation of the NARCOMS registry: diagnosis. Mult Scler. 2007 Feb 9;13(6):770–5.

10. UK MS Register. UK MS Regsiter Terms and Conditions [Internet]. UK MS Register, terms and conditions. 2012 [cited 2016 Nov 16]. Available from: https://www.ukmsregister.org/Portal/TAndC

11. Zigmond, A, Snaith, R.P. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scaninavica. 1983;67(6):361–70.

12. Brooks R. EuroQol: the current state of play. Health Policy. 1996 Jul;37(1):53–72.

13. Hobart J. The Multiple Sclerosis Impact Scale (MSIS-29): A new patient-based outcome measure. Brain. 2001 May 1;124(5):962–73.

14. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. Ann Neurol. 2011 Feb;69(2):292–302.

15. R Development Team. R: a language and environment for statistical computing [Internet]. 2011. Available from: https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing

16. Ford DV, Jones KH, Middleton RM, Lockhart-Jones H, Maramba ID, Noble GJ, et al. The feasibility of collecting information from people with Multiple Sclerosis for the UK MS Register via a web portal: characterising a cohort of people with MS. BMC Med Inform Decis Mak. 2012;12(1):73.

17. ISD Scotland. The Scottish MS Register [Internet]. The Scottish MS Register. 2016 [cited 2016 Nov 16]. Available from: http://www.msr.scot.nhs.uk/index.html

18. Confavreux C. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. Brain. 2003 Apr 1;126(4):770–82.
19. Boiko A, Vorobeychik G, Paty D, Devonshire V, Sadovnick D, the UBC MS Clinic Neurologists. Early onset multiple sclerosis: A longitudinal study. Neurology. 2002 Oct 8;59(7):1006–10.
20. Sumelahti M-L, Holmberg MHA, Murtonen A, Huhtala H, Elovaara I. Increasing Incidence in Relapsing-Remitting MS and High Rates among Young Women in Finland: A Thirty-Year Follow-Up. Mult Scler Int. 2014;2014:1–8.
21. MSIF. Atlas of MS [Internet]. Atlas of MS. 2013 [cited 2016 Nov 7]. Available from: https://www.msif.org/about-us/advocacy/atlas/
22. Paulus A, Hussack S, Kugler J. Multiple Sklerose – Ergebnisse einer Befragung der Mitglieder des DMSG-Landesverbandes e.V. in Sachsen-Anhalt: Durch welche Faktoren wird die Diagnosedauer beeinflusst? Fortschritte Neurol · Psychiatr. 2016 Aug 29;84(8):487–93.
23. Alonso A, Hernan MA. Temporal trends in the incidence of multiple sclerosis: A systematic review. Neurology. 2008 Jul 8;71(2):129–35.
24. Sellner J, Kraus J, Awad A, Milo R, Hemmer B, Stüve O. The increasing incidence and prevalence of female multiple sclerosis—A critical analysis of potential environmental factors. Autoimmun Rev. 2011 Jun;10(8):495–502.
25. Koch-Henriksen N, Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol. 2010 May;9(5):520–32.
26. Miller DH, Leary SM. Primary-progressive multiple sclerosis. Lancet Neurol. 2007 Oct;6(10):903–12.
27. Ribbons KA, McElduff P, Boz C, Trojano M, Izquierdo G, Duquette P, et al. Male Sex Is Independently Associated with Faster Disability Accumulation in Relapse-Onset MS but Not in Primary Progressive MS. Aktas O, editor. PLOS ONE. 2015 Jun 5;10(6):e0122686.
28. Waterson P. Health information technology and sociotechnical systems: A progress report on recent developments within the UK National Health Service (NHS). Appl Ergon. 2014 Mar;45(2):150–61.
29. Health and Social Care Information Centre. NHS ICD-10 5th Edition XML data files releases [Internet]. NHS Classifications ICD10 5th Edition. 2015 [cited 2016 Nov 10]. Available from: https://isd.hscic.gov.uk/trud3/user/guest/group/0/pack/28/subpack/259/releases
30. NHS Digital. UK Read Code [Internet]. UK Read Code. [cited 2016 Nov 10]. Available from: https://data.gov.uk/dataset/uk-read-code/resource/3314d8a4-ef2c-4a00-a7d7-ddec14b88519
31. Cecil Prescott. Internet users in the UK: 2016 [Internet]. Statistical Bulletin. 2016 [cited 2017 Apr 26]. Available from: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2016
32. Fox S, Purcell, Kristen. Chronic Disease and the Internet [Internet]. Pew Research Center. 2010 [cited 2016 Nov 9]. Available from:

http://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/
33. ONS. Population Estimates for UK, England and Wales, Scotland and Northern Ireland: mid-2015 [Internet]. Statistical Bulletin. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2015#uk-population-reached-651-million-in-mid-2015
34. Kremenchutzky M. The natural history of multiple sclerosis: a geographically based study 9: Observations on the progressive phase of the disease. Brain. 2006 Jan 6;129(3):584–94.
35. Leray E, Moreau T, Fromont A, Edan G. Epidemiology of multiple sclerosis. Rev Neurol (Paris). 2016 Jan;172(1):3–13.

Appendix 1

List of clinical sites that contributed data to this study

ABMU NHS Trust Morriston Hospital
Nottingham Hospitals NHS Trust, Queens Medical Centre
UCL London Hospitals NHS Trust National Hospital for Neurology and Neurosurgery
Belfast Health and Social Care Trust, Belfast City Hospital
Barts Health NHS Trust, Royal London Hospital
Shrewsbury and Telford NHS Trust, Royal Shrewsbury Hospital
Basildon and Thurrock University Hospitals NHS Foundation Trust, Basildon Hospital
University Hospital Southampton NHS Foundation Trust. Southampton General Hospital
Oxford University Hospitals NHS Foundation Trust John Radcliffe Hospital
Poole Hospital NHS Foundation Trust, Poole Hospital
Brighton and Sussex University Hospitals NHS Trust , Princess Royal Hospital
Royal Free London NHS Foundation Trust, Royal Free Hospital
Hampshire Hospitals NHS Foundation Trust, Basingstoke and North Hampshire Hospital
Salford Royal NHS Foundation Trust, Salford Royal Hospital
Northampton NHS Trust, Northampton General Hospital
Royal Devon and Exeter NHS Foundation Trust, Royal Devon and Exeter Hospital
Luton and Dunstable Hospital NHS Foundation Trust, Luton Hospital
Frimley Park NHS Foundation Trust, Frimley Park Hospital
Barking, Havering And Redbridge Hospitals NHS Trust, Queens Hospital
Mid Yorkshire Hospitals NHS Trust, Pinderfields General Hospital
Royal Cornwall Hospitals NHS Trust, Royal Cornwall Hospitals NHS Trust
Southend University Hospital NHS Foundation Trust, Southend Hospital

Appendix 2 UK MS Register Minimum Dataset
IP CAG

MS REGISTER

Minimum Data Set
Version 6.3

Name & Address

Date	/ /

Study ID						

TIER ONE (all fields to be completed)

MS Type now			
RR	SP	PP	Other
Conversion to SP (if applicable)			

Date of Diagnosis	/ /	**MS Type at Diagnosis**	RR SP PP Other
Date of onset	/ /		

No. of Relapses (RR only) (since last visit/year)		**Severity:** (circle)	Mild	Moderate	Severe

PAST Disease modifying Treatment (circle)					**Date Started**	/ /
Alemtuzumab	Avonex	Betaferon	Copaxone	Extavia	Fingolimod	Mitoxantrone
Natalizumab	Ocrelizumab	Rebif	Tecfidera	Teriflunomide	None	
Date Stopped	/ /	**Reason** (circle)	Lack of efficacy / Side Effects / Other			

PRESENT Disease modifying Treatment (circle)					**Date Started**	/ /
Alemtuzumab	Avonex	Betaferon	Copaxone	Extavia	Fingolimod	Mitoxantrone
Natalizumab	Ocrelizumab	Rebif	Tecfidera	Teriflunomide	None	

Current EDSS Score (1-10)		**Date EDSS Taken**	/ /	
Walking range : time and distance in meters	Time = _____ M = _____	Self-estimated	Trundle wheel	Treadmill

TIER TWO (to be completed IF patient is unlikely to register online)

Patient is: (tick)

Pregnant	Y	N

Smoker	Y	N
No. Per Day		
Smoked since:		

Onset Localisation (circle)	
Spinal	Cortex
Visual	Cerebellar/brainstem

Onset Symptoms (circle)		
Vision	Motor	Sensory
Coordination	Bowel/Bladder	Fatigue
Cognitive	Encephalopathy	Other

Person completing form: _____