

Lessons from “Lessons from Failure with the UK Web Archive” – The MMR Crisis, 1998-2010

At the previous RESAW conference, I presented a paper as part of a panel alongside Richard Deswarte and Rowan Aust about our work on the Big UK Domain Data for the Arts and Humanities project (BUDDAH).¹ My project had investigated the presence of British disability charities on the early web and how their sites might reflect on the changing nature of disability activism following the Disability Discrimination Act 1995.² The subtitle of that paper was ‘lessons from failure with the UK web archive’ – an admission of frustration with the limitation of the author as much as (if not more than) the accessibility of the material itself. As part of my current research project at the London School of Hygiene & Tropical Medicine on the history of vaccination since the Second World War, I have once again had cause to use the UK web archive for research. This paper outlines how this research has progressed and what lessons were learned from ‘lessons from failure’. My research questions were more focused and achievable, even if this may have made them less ambitious. One key conclusion, however, remains. British history from the late 1990s onwards will increasingly require knowledge of and disciplinary command over internet and web source material.

I should begin by saying – or admitting – that I do not consider myself a “historian of the internet”. I am currently writing a book on the history of British vaccination policy since the Second World War.³ My Masters and PhD are in the history of medicine, and my current position focuses on the history of public health. I am a historian of post-war Britain, and as such rely heavily on documentary source materials for my work on public health policy and the ways in which this was interpreted, accepted or resisted by the general public. I also have experience with oral history interviews.

My work, therefore, deals primarily with documents produced by central bodies, such as government departments, private companies, non-governmental organisations and so forth. These

¹ The event details and programme are available at <http://resaw.eu/events/international-conference-aarhus-june-2015/> (accessed 22 May 2017). For BUDDAH, see: <https://buddah.projects.history.ac.uk/> (accessed 23 May 2017).

² Gareth Millward, ‘Naïve researchers do bad history – lessons from failure with the UK web archive’, presented at RESAW, Web Archives as Scholarly Sources: Issues, Practices and Perspectives, 10 June 2015. See also, Gareth Millward, Peter Webster and Richard Deswarte. ‘Interrogating the archived UK web: Historians’ and Social Scientists’ Research Experiences’, presented at IHR Digital History Seminar Series, 4 November 2014. <https://ihrdighist.blogs.sas.ac.uk/2014/10/28/tuesday-4-november-interrogating-the-archived-uk-web-historians-and-social-scientists-research-experiences/> (accessed 22 May 2017).

³ See: <http://placingthepublic.lshtm.ac.uk/> (accessed 22 May 2017).

form the backbone of my work, but my interest in the public means that I also spend a great deal of time with newspapers, journals, pamphlets and other expressions of what might broadly be termed “public opinion”.⁴ For the period before the 1990s, these sources are relatively easy to find in well-catalogued repositories such as The National Archives or the British Library; and, increasingly, in digitized formats in databases such as the Millbank Hansard, the Times Digital Archive, JSTOR and so on.⁵ After the 1990s, many of these types of source are “born digital”. The rise of websites and the convention of publishing key press releases, reports and health educational materials online mean that archived websites become important repositories of information in their own right.⁶ Even for government departments, not all of these documents have a print equivalent stored in a library or traditional archive. Even more importantly, if we wish to explore “public opinion”, blogs, social media sites and personal webpages are critical. We can see not only individual’s responses to particular matters, but also get access to the webpages that influenced wider public opinions and attitudes.

What we learned from BUDDAH is that searching the archived web is a difficult task. The Wayback Machine is a much more useful tool when you know the URL of the site you are attempting to find.⁷ Trying to select anything approaching a representative sample on a topic as broad as “vaccination” is almost impossible, even using the sophisticated search options available through the British Library’s “Shine” interface.⁸ A full project investigating the MMR crisis on the internet is definitely worth doing, and if I had the time and resources in my current role I would certainly attempt it. However, the crisis itself accounts for less than a chapter in the monograph I am currently writing. At the same time, we know from other documentary accounts and memories of the crisis that the internet is generally considered to have played a major role both in the late nineties and in present day hesitancy surrounding vaccination. The government set up its own pro-MMR website to tackle what it saw as misinformation. The rulings of the General Medical Council on the conduct of key doctors was published online; so too was the government’s advice to GPs on immunization science and policy. As a historian living within the practical limitations of a three-year contract and my publisher’s demands, therefore, I am in somewhat of a conundrum. I *could* write a history of MMR solely from, for want of a better phrase, “traditional” documents. The medical and national press, memoirs, Hansard and official publications give more detail than could ever be processed by a single

⁴ On the use of the media as a source, see: Adrian Bingham, ‘Ignoring the First Draft of History?’, *Media History* 18, no. 3–4 (1 August 2012): 311–26, doi:10.1080/13688804.2012.721644.

⁵ Although these can be problematic in themselves. See: Tim Hitchcock, ‘Confronting the Digital: Or How Academic History Writing Lost the Plot’, *Cultural and Social History* 10, no. 1 (1 March 2013): 9–23.

⁶ For a history of medicine example of this, see: Martin Gorsky, ‘Into the Dark Domain: The UK Web Archive as a Source for the Contemporary History of Public Health’, *Social History of Medicine* 28, no. 3 (August 2015): 596–616, doi:10.1093/shm/hkv028.

⁷ <https://archive.org/> (accessed 23 May 2017).

⁸ <https://www.webarchive.org.uk/shine> (accessed 23 May 2017).

person. I would also be able to get at public reactions to the crisis through these media, and produce a narrative of government policy on MMR and how it changed over time. But while this is *possible*, it is somewhat lacking. If we know that activity on the internet was important, documents from the web need to be included and placed within the context of the wider media consumed by contemporaries.

My compromise was to follow up on URLs quoted in my other source materials and build these into my analysis. If URLs were provided, this gave some sort of indication about the importance of those links, or at least gave interesting illustrative examples of certain aspects of the MMR crisis. From there, the content of the sites could be brought in as evidence of activity by individuals and institutions. It was possible to triangulate information on the archived web with other verifiable source materials so that each document could add context to the others. It also meant I would analyse a smaller, concentrated sample of websites, giving more time to explore those sites in depth. Yet it was also methodologically problematic. Without reliable search methodologies, there may well have been either individual sites with a wealth of information and significance, or many smaller sources pointing to specific trends in the public's perception of MMR and vaccination in general. While not entirely necessary to produce the sort of history I intended to write, it does mean that there is a bias in the source material towards "establishment" and "official" narratives. Whether that fatally undermines the points I am trying to make will be the judgement of my referees and my peers. In any case, future histories of MMR will need to engage with this if they are to produce a balanced and richer version of events from multiple viewpoints.

Despite this "compromise", there is a wealth of accessible, relevant and useful information. A number of documents cite URLs, and one of the first things I noticed was how many of these links were broken. For smaller organizations or personal websites this was not particularly surprising. We know that "link rot" spreads quickly in databases (in 2014, 70 per cent of URLs in the Harvard Law Review and other journals no longer pointed to the cited information).⁹ Even government departments were not immune. The British Department of Health, for example, switched its main domain from doh.gov.uk to dh.gov.uk in 2004, completely restructuring the site in the process and removing a number of pages. In other cases, press releases were not necessarily archived and stored on the live web as a matter of course. For more official documents, this was not a major issue, since they were stored both in the publicly accessible Wayback Machine and through The National Archives' own snapshots of government websites.¹⁰ Individual and smaller websites could sometimes pose a

⁹ Jonathan Zittrain, Kendra Albert, and Lawrence Lessig, 'Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations', *Legal Information Management* 14, no. 2 (June 2014): 88–99, doi:10.1017/S1472669614000255.

¹⁰ <http://nationalarchives.gov.uk/webarchive/> (accessed 23 May 2017).

larger problem, as will be discussed later in this paper, but were for the most part accessible *enough* to be used as important sources in their own right.

The MMR crisis on the web

The MMR crisis was a period of around six years from the publication of a paper in *The Lancet* and subsequent press conference by Andrew Wakefield and other researchers and clinicians at the Royal Free Hospital.¹¹ It was claimed that there may have been a link between a particular type of autism and the MMR vaccine. While the scientific legitimacy of this claim was immediately and thoroughly refuted, the vocal minority of dissenting voices gained press attention in the wake of a number of scandals that had dented the reputation of the medical profession in the UK.¹² This peaked in the early months of 2002 when a BBC documentary repeated the doubts about MMR's safety.¹³ Prime Minister Tony Blair refused to say whether his young son, Leo, had received the vaccine, creating a media storm that he later regretted.¹⁴ It was, however, short-lived.¹⁵ In 2004, a Channel 4 documentary and *Sunday Times* exposé series exposed Wakefield's unethical clinical practices and conflicts of interest which put an end to the crisis both in the popular and medical press.¹⁶ During this period, however, the percentage of children in England receiving their first dose of MMR by their second birthdays dropped from 90.8 in 1998 to 79.9 in 2004.¹⁷ It has become an archetypical example for public health professionals explaining where and how vaccine scares can occur, and many use it as a "lesson from the past" to help present-day practice.¹⁸ For these reasons, it is important that we have the tools to

¹¹ AJ Wakefield et al., 'RETRACTED: Ileal-Lymphoid-Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children', *The Lancet* 351, no. 9103 (28 February 1998): 637–41, doi:10.1016/S0140-6736(97)11096-0.

¹² The two highest profile cases being the BSE/vCJD crisis and the contaminated blood scandal.

¹³ Abi Berger, 'Dispatches. MMR: What They Didn't Tell You', *BMJ* 329, no. 7477 (25 November 2004): 1293, doi:10.1136/bmj.329.7477.1293.

¹⁴ Tony Blair, *A Journey* (London: Hutchinson, 2010), 646.

¹⁵ Tammy Speers and Justin Lewis, 'Journalists and Jabs: Media Coverage of the MMR Vaccine', *Communication & Medicine* 1, no. 2 (2004): 171–81, doi:10.1515/come.2004.1.2.171.

¹⁶ Brian Deer, 'How the Case against the MMR Vaccine Was Fixed', *BMJ* 342 (6 January 2011): c5347, doi:10.1136/bmj.c5347. Wakefield's co-authors retracted the conclusion to the article in 2004, and the article itself was retracted by the *Lancet* editorial board in 2010. Simon H Murch et al., 'Retraction of an Interpretation', *The Lancet* 363, no. 9411 (6 March 2004): 750, doi:10.1016/S0140-6736(04)15715-2; The Editors of The Lancet, 'Retraction—Ileal-Lymphoid-Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children', *The Lancet* 375, no. 9713 (February 2010): 445, doi:10.1016/S0140-6736(10)60175-4.

¹⁷ It did not exceed the 1998 levels until 2012. See: NHS Digital, 'NHS Immunisation Statistics: England, 2015-16' (London: National Statistics, 22 September 2016), <http://content.digital.nhs.uk/catalogue/PUB21651>; Health and Social Care Information Centre, 'NHS Immunisation Statistics: England, 2004-05' (London: National Statistics, 22 September 2005), <http://content.digital.nhs.uk/catalogue/PUB00176>.

¹⁸ Heidi J. Larson et al., 'Understanding Vaccine Hesitancy around Vaccines and Vaccination from a Global Perspective: A Systematic Review of Published Literature, 2007-2012', *Vaccine* 32, no. 19 (17 April 2014): 2150–59, doi:10.1016/j.vaccine.2014.01.081. On the use of "lessons from the past" in public health (and the

critically evaluate the crisis – and part of this will require us to have the ability to historically analyse websites from the period.

For the purposes of this paper, I will show three key examples of websites found in the archives that are methodologically and historically interesting. Perhaps the most obvious and important use of the Wayback Machine was in tracking changes to the government’s website which had been established to counter claims that MMR was not safe. Providing help and advice to parents, it had a section which answered questions submitted from the general public. I identified the website from an article in the *British Medical Journal*, and found that the URL mnrthefacts.nhs.uk no longer worked.¹⁹ The archive therefore proved useful not just for finding the website but in showing how it evolved as the “news” surrounding the science on MMR and the content of the “frequently asked questions” changed over time. Thus, while the mere existence of mnrthefacts.nhs.uk tells us much about the government’s attitudes towards the public and health communication, the archive helps us to show how websites are not static documents like reports, leaflets or press releases. The site served many of the functions of these static technologies, but responded to public input (within its own set parameters). The site’s final copyright notice reads 2004. Moreover, while the questions and news sections continue to be maintained until the late 2000s, the core content of the site changes little after this point.²⁰ When we triangulate this information with what we know from other printed and digital sources, we add weight to the claim that the crisis effectively ended with the “exposure” of Wakefield in 2004. For example, the number of mentions of “MMR” in the ProQuest UK Newsstand database for 2005 is barely half the number for 2004.²¹

The second key page was a report written by an anti-vaccination campaigner in the United States and hosted on a university server. It came to my attention when a public health worker cited it in a letter to the *British Medical Journal*, arguing that it had created a lot of discussion among his or her non-medical friends.²² While we cannot be sure from this one citation what impact this page had on the general public, it acts as an interesting case of how the early web was used to convey

limitations of such approaches), see: Virginia Berridge, Martin Gorsky, and Alex Mold, *Public Health in History*, 1 edition (Maidenhead: Open University Press, 2011).

¹⁹ Mirza Muminovic, ‘Mmr’, *BMJ* 325, no. 7364 (14 September 2002): 604, doi:10.1136/bmj.325.7364.604/a.

²⁰ See various captures in the Wayback Machine, https://web.archive.org/web/*/http://www.mnrthefacts.nhs.uk:80/ (archived between 8 September 2002 and 22 February 2010. Accessed 22 May 2017).

²¹ There were 1200 results for 2004, 620 for 2005. Search for string “MMR” in ProQuest UK Newsstand on 11 December 2016, 13:50. Accessed through Senate House Library, University of London. See: <http://proquest.libguides.com/UKNewsstand>. Total results for 1995-2015, n=19,904.

²² British Medical Journal Publishing Group, ‘MMR Vaccination and Autism 1998’, *BMJ* 316, no. 7147 (13 June 1998): 1824, doi:10.1136/bmj.316.7147.1824.

unconventional health messages.²³ In the letter, the health worker argues that the page is ‘written in the style of a medical journal, which lends it more authority than it may merit’, and it is indeed presented as a matter-of-fact “report”. This choice of presentation may be useful to historians in itself. We know, for example, that anti-vaccination groups have consistently delivered their information as rational and scientific, borrowing the styles and fashions from the groups whose authority they sought to undermine.²⁴ It also tells us something of the imagined concerns of public health workers, even if there may be many other factors in parents’ decision-making surrounding vaccination. Regardless, like mmrthefacts.nhs.uk, historians can place this page into context through triangulation with other sources. The author, Alan G. Phillips continued to campaign against vaccinations, and currently has a law firm specialising in filing vaccine exemptions with various bodies in the United States. On his personal site, he has a Curriculum Vitae – which shows that he was working at the University of North Carolina (the domain on which the page is hosted) at the time the letter was published in the *British Medical Journal*.²⁵ We also know that other public health professionals were worried about the impact of the internet, and they expressed this in print.²⁶ If Phillips’ webpage in itself may simply be one example of many such attempts to weaken faith in vaccination, it is at least useful enough as a source (in the wider context of our evidence base) to draw some meaningful conclusions.

The third example relies on more than just the Wayback Machine. A public health worker, A Rouse, highlighted the work of the Society for the Autistically Handicapped. Their website, he claimed, showed that Andrew Wakefield’s *Lancet* paper was built on a conflict of interest. It asked serious questions about how Wakefield found his patients for his work, and suggested that many of his children came via a legal firm constructing a case against the government for compensation for injuries caused, allegedly, by MMR.²⁷ We know from the General Medical Council’s judgement on Wakefield and the work of Brian Deer that this was indeed the case.²⁸ However, the link provided by Rouse was

²³ Alan Phillips, *Dispelling Vaccination Myths*, last revision February 1997.

<https://web.archive.org/web/19990503175555/http://www.unc.edu:80/~aphillip/www/vaccine/dvm.txt> (accessed 16 May 2017, archived 3 May 1999).

²⁴ For this long history, see: Nadja Durbach, *Bodily Matters: The Anti-Vaccination Movement in England, 1853–1907* (Durham, NC: Duke University Press, 2005).

²⁵ Curriculum Vitae, August 2016. <http://www.vaccinerights.com/pdf/2016AlanPhillipsCurriculumVitae.pdf> (accessed 16 May 2017).

²⁶ British Medical Journal Publishing Group, ‘Health Professionals’ Attitudes to MMR Vaccine’, *BMJ* 322, no. 7294 (5 May 2001): 1120, doi:10.1136/bmj.322.7294.1120; Richard Smith, ‘Do Patients Need to Read Research?’, *BMJ* 326, no. 7402 (12 June 2003): 1307, doi:10.1136/bmj.326.7402.1307; Ian Hargreaves, Justin Lewis, and Tammy Speers, ‘Towards a Better Map: Science, the Public and the Media’ (London: Economic and Social Research Council, 2003), 45.

²⁷ A Rouse, ‘Autism, Inflammatory Bowel Disease, and MMR Vaccine’, *The Lancet* 351, no. 9112 (2 May 1998): 1356, doi:10.1016/S0140-6736(05)79082-6.

²⁸ Deer, ‘How the Case against the MMR Vaccine Was Fixed’. General Medical Council, ‘Dr Andrew Jeremy Wakefield, Determination on Serious Professional Misconduct and sanction’ (24 May 2010). <https://web.archive.org/web/20120320005240/http://www.gmc->

not to be found anywhere in the Wayback Machine. Given that the printed URL was on the .uk webspace, I searched the British Library's internet archive using the Shine interface, using the string "Society for the Autistically Handicapped", and limiting the results to 1999 (the nearest year to Rouse's letter for which there were any results).²⁹ This showed that there was a missing "r" in the printed URL – and I was able to access the Society's page from the search results.³⁰ The lack of snapshots in the Wayback Machine suggests that this was not a heavily linked-to site, but after following various hyperlinks I was able to locate a version of the page Rouse references after the site changed its name and domain.³¹ One of the report's authors was Richard Barr, a lawyer in the Legal Aid funded case against the government on behalf of autistic parents.³² Once again, when we triangulate the page with other documents we get a sense of what may have been known about Wakefield's activities before his exposure in 2004. Though much deeper research would be necessary to ascertain the degree to which this particular website (or any others) actually influenced public opinion or reflected a representative view of the debate, it does give us a sense of the sorts of activities being pursued online by vaccine sceptic groups and parents concerned with the MMR-autism link.

Conclusions

In my work and through the examples I have provided, internet archival material is an addition to my existing source base. It adds context, colour and nuance, giving greater weight to arguments or problematizing some of the claims made by actors in the policy-making process. A more social or cultural history of MMR specifically would almost certainly need to go further and engage more with search and discovery methodologies. However, as with all forms of history a single source base is not enough to write a balanced, detailed account. "Traditional" documentary materials still matter, and are in themselves a vital tool in the discovery process; as well as adding necessary context to the provenance and purpose of websites and their authors.

Without the experience of BUDDAH, I would probably not have used these sources or the internet archive in any meaningful way. I may have uncritically used the Wayback Machine to look at

[uk.org/Wakefield_SPM_and_SANCTION.pdf_32595267.pdf](http://www.uk.org/Wakefield_SPM_and_SANCTION.pdf_32595267.pdf) (captured 20 March 2012, accessed 10 November 2016).

²⁹ See:

https://www.webarchive.org.uk/shine/search?query=Society+for+the+Autistically+Handicapped&tab=results&action=search&facet.in.crawl_year=%221999%22 (accessed 22 May 2017).

³⁰ *The Lancet* printed "www.mplc.co.uk"; the actual site was hosted at "www.rmplc.co.uk".

³¹ Richard Barr and Kirsten Lamb, Vaccines Fact Sheet, updated 3 June 1997.

<http://web.archive.org/web/20010405224319/http://www.autismuk.com:80/index1sub4.htm> (archived 5 April 2001, accessed 16 May 2017).

³² Richard Barr, 'Autism, Bowel Inflammation, and Measles', *The Lancet* 359, no. 9323 (15 June 2002): 2112, doi:10.1016/S0140-6736(02)08917-1.

a couple of snapshots, but I would not have had the knowledge of Shine or the context of webcrawling and the limits of web archive technology to guide my studies. Does this reflect my own ignorance, or is it typical of “documentary historians” in my field more widely? If the latter, then we may have a problem. Even in a document-heavy research project such as this one, the internet archive proved invaluable in establishing the timeline of events in the MMR crisis. I have not gone into detail about it here, but various press releases and reports from academic, governmental and non-governmental bodies were only available (or, at least, most easily accessible) through web archives.³³ These are not simply going to be a nice addition to our histories of the 1990s and 2000s – they will be integral to good historical practice. It is important, therefore, that historians are both aware of the existence of web archives and the limits of what they can do with them. And this applies not just to technologically and methodologically innovative projects involving link analyses, text mining and “big data” techniques – “traditional” historians will have to get their heads around this, too. Hopefully, the research outputs from conferences such as this will begin to spread that message.

³³ See the ESRC’s report on media coverage on MMR: Hargreaves, Lewis, and Speers, ‘Towards a Better Map’. http://web.archive.org/web/20110526211506/http://www.esrc.ac.uk/images/towards_a_better_map_tcm8-13558.pdf (archived 26 May 2011, accessed 23 May 2017); the Medicines Controls Agency, ‘Report of the working party on MMR vaccine’ (London: Medicines Control Agency, 1999). <https://web.archive.org/web/20030210185232/http://www.mca.gov.uk/inforesources/publications/mmr1.htm> (archived 10 February 2003, accessed 23 May 2017); General Medical Council, ‘Dr Andrew Jeremy Wakefield’.