TURNING PAPER INTO GOLD
Measuring Doctors' Readership in a Cost Conscious Market

Richard Asquith, BMRB, KMR Group (UK)

Introduction

The National Medical Readership Survey (NMRS) is a study of doctors' reading habits in the UK, sponsored by an industry committee of interested parties – the Joint Industry Committee of Medical Advertisers for Readership Surveys (JICMARS). The survey provides data that are used to plan and trade advertising space in medical journals which are distributed, largely free of charge, to doctors in general practice (GPs).

Since the survey's inception in the late 1970s, data had been collected via face-to-face interviews amongst a pre-selected random sample of the GP population. However, the pharmaceutical advertising market has been under significant pressure in recent years due to changes in government policy reducing GPs' freedom to prescribe branded drugs. Between 1999 and 2006 the number of advertising pages in journals approximately halved and this, coupled with increasing survey costs as response rates gradually declined, meant that the study in its original form had become unviably expensive.

Consequently, JICMARS was forced to consider alternative, lower cost methodologies, but critically it had to retain the Gold Standard status and unqualified industry acceptance of the survey in the marketplace. Prior to the formation of JICMARS publishers conducted their own research of varying quality resulting in a free-for-all of readership claims and counter claims. Loss of the industry currency could have heralded a return to this uncertainty and potentially damaged the advertising market further.

This paper outlines how an extensive programme of testing allowed JICMARS to introduce a new lower cost methodology for the NMRS, whilst avoiding advertising market anarchy.

The Face-to-Face Interview Methodology

In summary, the original survey methodology employed the following approach:

- Several commercial mailing lists operate in the market, which aim to list all practicing GPs. The selected list was first stratified by region and then all *fully-qualified* GPs were grouped within practice, ordered by descending size of practice.

- A random sample of GPs was drawn from the list using a fixed sampling interval.

- Practices containing a sampled GP were phoned and the list of all practicing GPs including trainees (Registrars) checked against the original list. Registrars were identified in this way as they spend relatively short periods of time at a practice and the mailing list was known to be least accurate for this group of GPs. The screening call also identified and removed part-time and temporary GPs (locums) at the practice.

- The Marchant-Blyth technique was used to sample those GPs, including Registrars, who had been added to the original list at the screening stage, in their correct proportions.

- The sampled GPs were sent letters of introduction and then phoned and where possible appointments made to visit them for a 20 minute interview at their practice or home.

- The appointments were issued to local interviewers who visited the practice and conducted the CAPI interview.

- The CAPI interview sequence established for each publication – ever read, last read in the issue period, first time readership of older issues, frequency of reading, depth of reading and what done with the publication once read, followed by a wide range of classification questions.

- The respondent incentives were adjusted over time to try to maintain high levels of co-operation. Latterly, a GP was given £30 as a "token of our appreciation" for participating in the research. This is substantially below standard GP consultation rates.

Until 1997, the response rate was maintained at above 60%, but over the past ten years this has steadily declined, increasing the cost of conducting the survey. For the past few years the response rate had fluctuated around 40%.

The NMRS standard measure of readership is "Total Average Issue Readership". The survey employs the recency approach for readership estimation, but a technique has been developed to identify and include "Additional New Readers". Additional New Readership is first time readership of older issues of a publication within the issue period when the current issue has also been read in that period – readership that would otherwise be lost using the standard recency methodology. This is judged to be particularly important for medical and other professional journals which are commonly stored for future reference or when time permits for reading.

Within the CAPI interview, Additional New Readers were identified using the following two questions (for a weekly publication):

> During the last seven days have you read, looked through or referred to only the latest issue of xxx; both the latest issue and an older issue; or only an older issue (but not the latest issue)?
>
> *If an older issue read:*
> Was one of those older issues one which you read, looked through or referred to for the first time in the last week?

Additional New Readers are added to Average Issue Readers (ie read any issue for the first time within the issue period) to create Total Average Issue Readership. Therefore, certain respondents are in practice being counted twice – once as an Average Issue Reader and again as an Additional New Reader.

Total Average Issue Readership (12 month rolling data)
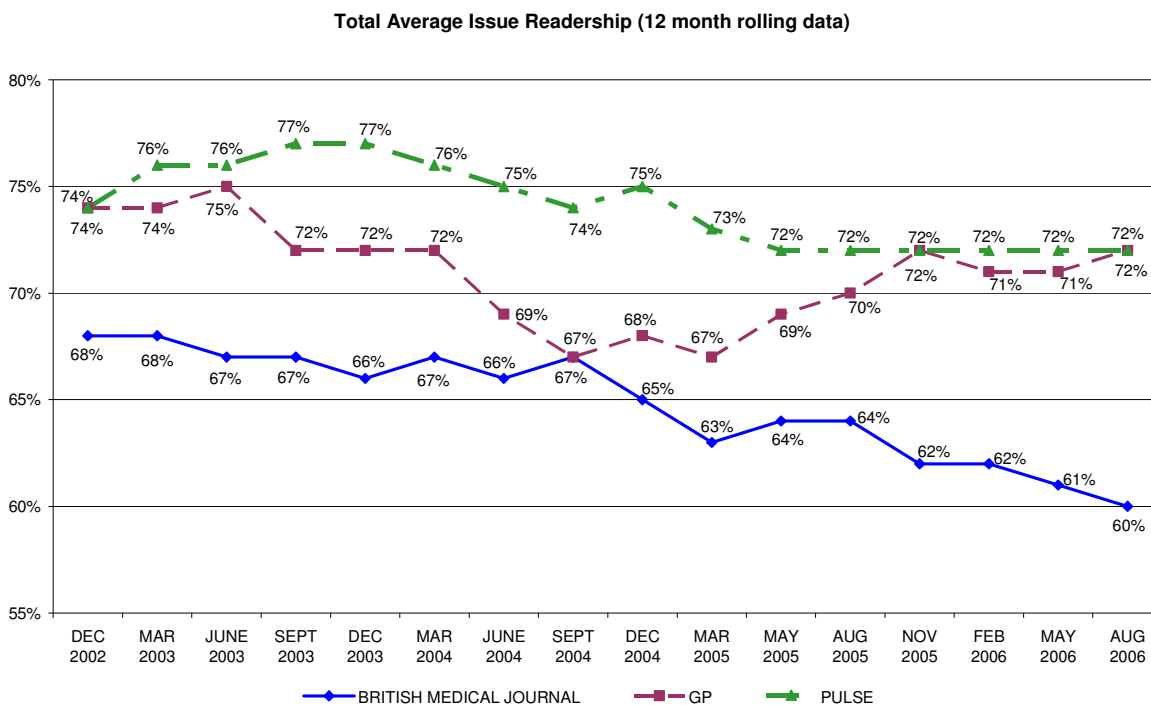
The annual sample size of 1,000 GPs (out of a total UK population of over 35,000) has yielded very consistent estimates of Total Average Issue Readership over the years. Clearly, this stability has been beneficial to the publishers and the pharmaceutical advertising industry generally and has contributed significantly to maintaining financial support for the study for many years.

Exploring Alternative Methods

Publishers provide the majority of the funding for the NMRS and the substantial decline in advertising revenue over the past few years prompted a review in 2005 to assess whether or not an alternative survey methodology could provide a more cost-effective solution. JICMARS initially considered CATI interviewing and a variety of self-completion methodologies – online, mailed PDAs and mailed paper questionnaires.

It is worth saying at this point that consideration was also given to dropping the initial telephone screening / list confirmation. However, it was decided that to dispense with this stage would have compromised the sample design too much at a time when a change to the data collection methodology could also introduce uncertainty in the market.

During the course of the evaluation of possible alternative methodologies, the CATI approach was rejected as it was felt that visual prompting of publications, several of which have very similar titles, was essential for the survey. Whilst it is of course possible to provide visual prompting of covers and mastheads at the same time as conducting a telephone interview, it was felt that pre-mailing could influence subsequent responses unduly and simultaneous online viewing would be impractical for busy (and heavily researched) GPs and damage response rates.

Self-completion of some sort was thought to be most viable as the GP could fit completion around their schedules. However, a significant reduction in the questionnaire length would be required to try to maintain an acceptable response rate.

At this stage it is fair to say that a decision was made to test a self-completion *paper* questionnaire based on the belief (rather than evidence) that it would provide the most cost-effective and practical solution for GPs. In 2005 85% of GPs had access to a PC with broadband internet connection and it was expected that this would continue to increase over the coming years, but all sides of the industry had serious doubts that a non-panellised sample of GPs could be persuaded to go online and seek out a readership questionnaire and thereby achieve an acceptable response rate. A number of commercially run GP panels exist in the UK and are widely used for research, but they were felt not to be sufficiently representative for the NMRS. Observation of the face-to-face interviews had shown conclusively that convenience for the respondent was absolutely critical in generating acceptable response rates. And if the response rate was very low, the online methodology was likely to skew the sample towards younger, more technologically-minded respondents who would have different reading habits.

Mailing out of PDAs was rejected on grounds of cost and, again, a fear that it might skew the sample. So the decision was taken to conduct a full-scale parallel run of the postal paper methodology (with initial telephone screening). We estimated that the change in methodology could halve the cost of the survey, but there was the possibility that a substantially lower response rate would produce data that the market could not accept. In 2005 the face-to-face survey was achieving response rates of 40-45% - we didn't believe that it would be possible to maintain that level of response with a postal paper questionnaire, but guestimates of what we would achieve ranged from 15-35%!

We chose to test a £20 incentive, paid on receipt of a completed questionnaire – somewhat lower than the £30 incentive we had paid for a CAPI interview, but not ungenerous for a self-completion exercise that would take the respondent 5-10 minutes to complete.

Piloting the Paper Questionnaire

Testing of the postal paper questionnaire was initially planned to take place over two stages:

1. A small scale pilot mailing out 35 questionnaires with telephone follow up to test GPs understanding of the questions and to get a feel for their preparedness to complete the questionnaire.

2. A full parallel run for one fieldwork quarter (c.250 interviews) to assess response rate, sample profiles and variability in the readership data.

Prior to this, however, it was necessary to agree cuts in the length of the questionnaire and following extensive debate this was reduced from a face-to-face interview lasting 15-20 minutes to a four page questionnaire covering readership (including some quality of reading metrics) and basic classification data.

The small scale pilot exercise, in which eight completed questionnaires were returned, reassured us that the postal paper methodology might work and the respondents claimed at least to understand the questions and to find the questionnaire easy to complete.

The First Parallel Run

Fieldwork for the first parallel run of the paper questionnaire alongside the CAPI survey took place from December 2005 to February 2006 (Quarter 1 2006) and the key sample profile findings were as follows:

Q1 2006	Paper	CAPI	GP Population
Sample Size	311	266	
Response Rate	28%	44%	
	%	%	%
Male	64	65	60
Female	36	35	40
Year of Qualification:			
Pre 1968	6	6	7
1969-78	24	27	23
1979-88	41	33	34
Post 1988	28	34	36
Region:			
Scotland	13	11	14
North	23	21	23
Midlands	19	19	19
South	46	49	44

As expected, the response rate for the paper questionnaire was lower than we had been achieving using a face-to-face methodology, but not as poor as some had feared. The sample profile across the main classification groups for the paper questionnaire was pretty good – the main divergence from the population being an apparent under-representation of more recently qualified GPs.

However, the acceptability of the new methodology would be primarily dictated by the consistency of the readership data. JICMARS does not as a matter of course publish data on individual fieldwork quarters, choosing to provide rolling 6 and 12 month data. We felt that comparisons of individual title's readership data from the paper and CAPI surveys would be too susceptible to sampling variation so titles were grouped by frequency of publication (and therefore largely by genre – weekly news magazine or monthly clinical journal) for analysis purposes.

Gross Readership Comparisons Q1 2006

	Weekly / Fortnightly Publications		Monthly Publications	
	Paper	**CAPI**	**Paper**	**CAPI**
Ever See	520	518	424	429
AIR	290	307	196	202
ANR	11	29	7	16
TAIR	301	335	203	218
Frequency of Reading:				
Every/most issues	302	327	108	146
About half	65	58	43	42
Occasional / v few	123	130	158	147
Total responses	490	515	309	335
Reading Intensity:				
Cover-to-cover/most	214	208	69	82
About half	108	101	68	69
Less / v little	144	206	155	185
Total responses	466	515	292	336

AIR = Average Issue Readership
ANR = Additional New Readers
TAIR = Total Average Issue Readership

The table shows gross readership comparisons, ie the sum of percentages for all of the titles within each of the two publication frequency groupings.

Whilst the measures of Ever See and Average Issue Readership levels were close for the two methodologies, the first version of the paper questionnaire produced significantly lower levels of Additional New Readership than the CAPI survey which in turn led to lower levels of Total Average Issue Readership, the standard measure of readership in the NMRS. Unsurprisingly, this made the findings somewhat less palatable to the publishers! Under-recording of ANR clearly had to be addressed before the new methodology would be accepted by the survey sponsors.

The Quality of Reading measures also produced some interesting comparisons. For both frequency of reading and intensity of reading the paper questionnaire captured fewer responses in total, the implication being that some respondents who claimed to read a title were not following the routing instructions correctly and not completing the subsequent questions. There was also a suggestion that the least committed readers of each publication (those who read occasional issues and/or very little of it on average) were more susceptible to skipping the quality of reading questions.

So our conclusions were that there were grounds for optimism, but if the paper questionnaire was to be an acceptable methodology, we had to increase the capturing of Additional New Readers, we must try to improve compliance with the routing to the quality of reading questions and, if possible, increase response rates especially amongst the most recently qualified GPs.

Additional New Readers

As previously discussed, identification of Additional New Readers to generate estimates of Total Average Issue Readership is thought to be crucially important for a survey of GPs readership. The susceptibility of the Recency model to miss parallel reading is deemed to be unacceptable in this market. However, the questioning process employed in a face-to-face interview involved routing which could be easily programmed into the CAPI script but was difficult to transfer to a fixed format paper questionnaire.

Our first attempt to replicate this in the paper questionnaire was:

> Please answer Q3 for every publication you crossed in the shaded area at Q2. At Q3 you may choose more than one option for each publication. In the last 7 days, did you read…?
>
> - The latest issue
> - An older issue for the first time
> - Re-read an older issue

Our efforts to simplify the question had clearly gone too far and as a consequence it was not prompting respondents either to recognise instances of reading older issues for the first time as well as the current issue or they were failing to multi-code their answers.

We therefore came to the view that the question had to be divided up to ensure that due consideration was given. For the second parallel run we used the following questions:

> Please answer Q3 and Q4 for every publication you crossed in the shaded area at Q2.
>
> Q3
> In the last 7 days, did you read the latest issue?
>
> - Yes
> - No
>
> Q4
> And in the last 7 days, did you read an older issue for the first time and/or re-read an older issue? If appropriate mark both options.
>
> - An older issue for the first time
> - Re-read an older issue
> - Neither

The Second Parallel Run

The second full scale pilot was conducted from March to May 2006. Once again, the sample profile was close to the universe and CAPI profiles. On this occasion, despite a slightly lower response rate than in the first parallel run, there was no shortfall of more recently qualified GPs.

Q2 2006	Paper	CAPI	GP Population
Sample Size	274	256	
Response Rate	26%	42	
	%	%	%
Male	57	59	60
Female	43	41	40
Year of Qualification:			
Pre 1968	4	7	7
1969-78	19	20	23
1979-88	34	33	34
Post 1988	42	41	36
Region:			
Scotland	13	11	14
North	23	23	23
Midlands	16	16	19
South	49	49	44

The revisions to the Additional New Reader questions had the desired effect of increasing claims, but on this occasion the base AIR data was slightly lower from the paper questionnaire so overall the TAIR figures remained below the equivalent CAPI data.

Gross Readership Comparisons Q2 2006

	Weekly / Fortnightly Publications		Monthly Publications	
	Paper	CAPI	Paper	CAPI
Ever See	539	513	470	424
AIR	280	297	178	198
ANR	24	19	16	14
TAIR	302	319	195	212
Frequency of Reading:				
Every/most issues	315	309	125	137
About half	74	62	48	54
Occasional / v few	102	141	123	137
Total responses	491	512	296	328
Reading Intensity:				
Cover-to-cover/most	194	179	86	75
About half	113	121	49	75
Less / v little	131	215	138	177
Total responses	438	515	273	327

AIR = Average Issue Readership
ANR = Additional New Readers
TAIR = Total Average Issue Readership

Interestingly, the screen-in "ever see" figures were above the equivalent CAPI data so any shortfall in AIR was not due to a lack of recognition of the titles in the new questionnaire format.

Improved routing instructions to the quality of reading questions increased the volume of responses and results for the highest level of engagement were similar to the CAPI data. However, some shortfall remained amongst readers who were less committed to the publications.

The analysis shown so far has compared gross readership within frequency groups, but we also examined the range of differences for individual titles.

% point differences in AIR – Paper v CAPI:

- Paper 7-10 points below CAPI   1 publication
- Paper 3-6 points below CAPI   5 publications
- Paper -2 to +2 points v CAPI   5 publications
- Paper 3-6 points above CAPI   1 publication
- Paper 7-10 points above CAPI   0 publications

Naturally, there was concern about the divergence of individual titles, but on balance it was decided that the overall shape of the paper questionnaire data was sufficiently close to the CAPI data for the survey sponsors to make the switch and this was implemented from the final fieldwork quarter of 2006. Since only 6 and 12 month rolling data are published, JICMARS had the reassurance of knowing that data from the paper questionnaire would be gradually introduced quarter by quarter which would help to smooth any shifts in figures for individual titles.

Going Live

The story should end here. We had fully tested the paper questionnaire methodology, knew that it would produce comparable data to the CAPI survey at substantially lower cost to the clients and we were aware of the scale and general direction of possible divergences.

And in almost every respect the readership recorded in that first quarter using the new methodology matched our and the market's expectations. The response rate was 32%, not so far off CAPI levels.

However, one publication recorded a TAIR which was significantly out of line from past data. The scale of the difference was such that we felt unable to publish its data without gaining a better understanding the cause. Two possible explanations were hypothesised:

1. This particular publication is different from all other titles on the survey as it is primarily a listing of drugs published monthly. This difference was not thought to present a problem for the CAPI interview but our fear was that something about its format and use by GPs affected measurement via the self-completion paper questionnaire. The monthly publication list includes a number of more specialist niche medical interest titles and there was thought to be a greater risk that respondents wouldn't look at the whole list of monthly titles carefully enough. Looking back at the pilot data there was evidence of divergence for this particular title but not to the extent that rang alarm bells at the time.

2. The publication had had a major revamp and relaunch at the start of the fieldwork quarter including the way in which the drugs are listed and categorised which could have affected use of the title and its readership levels. The publisher told us that qualitative and anecdotal reaction to the new format had been exceptionally positive, but we knew that it also represented a big change to the way the publication might be used by GPs.

Before we published any data for this title incorporating the new methodology further research was needed to establish which of the two hypotheses was correct, or perhaps it was a combination of the two.

The approach we took was to field two variants of the postal survey at the same time as the ongoing main fieldwork (whose sample was temporarily boosted to provide an adequate base for direct comparison). The first variant identified the publication more clearly as a prescribing directory alongside all of the other monthly journals. The second variant had a much shorter list of the most widely read publications – four weeklies and the one monthly title of concern. There was therefore minimal risk that the title might not receive careful consideration within this very short list of publications. Our hope was that the data produced by all three versions of the questionnaire, the standard questionnaire, the enhanced full version and the short version, would be sufficiently close to reassure us that a real change in reading behaviour was being picked up by the paper questionnaire post relaunch.

Much to our relief, this is what the results showed. It was concluded that the divergence for this publication was due at least primarily to the new format and relaunch. Confidence in the new postal paper methodology was restored.

Summary

Readership research is never immune from the economics of the markets being measured. In the case of the NMRS a substantial decline in advertising revenue meant that the CAPI methodology for the accepted gold standard survey had become unviably expensive. However, a return to uncontrolled proprietary readership studies funded by individual publishers was prevented by a carefully managed programme of ongoing consultation with the sponsors and extensive testing of a lower cost paper self-completion questionnaire, ensuring continued support for a joint industry survey from all sides of the industry.