Electronic Diaries

Applications and What Works in the Field

Michael R. Hufford and Alan L. Shields

A literature review reveals that better subject compliance and more accurate field data can result from providing subjects with electronic rather than paper diaries.

Electronic diaries, or EDs, are beginning to replace paper diaries as the method of choice for collecting self-report data from subjects in the field. Clinical researchers have used paper diaries for that purpose since the 1940s, but in an attempt to overcome subject noncompliance and data-quality problems associated with paper diaries, they now often use EDs. The application of EDs to clinical research is flourishing thanks to improvements in the reliability of ED hardware and software, the documented success of ED applications, and the release of regulatory guidance for the use of EDs in clinical research.

When subjects are the end users of electronic devices in clinical research, however, important issues arise about subject acceptance and compliance, and the regulatory integrity of these systems. This review examines the empirical, peer-reviewed literature on EDs to help the reader better understand the various applications of EDs to clinical research and the factors associated with the success and failure of ED use in the field.

Rationale for developing EDs

The application of EDs to the collection of subject self-report data in clinical research has its roots in three persistent problems associated with paper diaries:

- poor data quality
- long delays to data lock
- poor subject compliance.

Data quality. Paper diaries often produce data of poor quality. Specifically, paper diary data are often hard to read and contain missing or out-of-range data. Furthermore, subjects often enter extraneous information that must then be transcribed, coded, and entered into a database. Because of these problems, paper diary data can generate an extraordinary number of queries. The extensive data cleaning required to use paper diary data creates a significant burden in time and money. This results in paper diary data that is difficult to use and has impeded the use of diary methods in clinical research.

Data lock. The large number of queries and generally poor data quality result in long delays between last subject out and data lock. Studies that have examined the difference between locking paper and electronic data typically report that the EDs reduced time to data lock by approximately 80%. The significant differences in time to data lock have helped to show that robust, field-tested EDs can not only be a scientific success, but also can produce meaningful cost savings for the sponsor of a clinical trial.

Subject compliance. Subjects are frequently noncompliant with paper diaries. Investigators and clinical research assistants involved in conducting paper-diary trials find that many subjects exhibit “parking lot compliance”—hoarding diary cards and completing them in batches. A study by Litt and colleagues, for example, revealed that fully 70% of subjects falsified at least some of their diary entries every day by writing the time they
The search was limited to publica-
diaries, electronic diary
ject experience.7
is also likely that back-filled entries do not
While, diaries in the first place. It
is also likely that back-filled entries do not
vated the use of diaries in the first place. It
spective memory processes that moti-
faked diary entries rely on the very retro-

This literature review of EDs in clinical
research is based on a search of the
Medline (www.ncbi.nlm.nih.gov/pubmed) and the
American Psychological Association’s
PsychInfo database (member only, but
abstracts can be viewed at www.apa.org/
psycinfo/). The search terms used for this
review were computer diary and computer
diaries, computerized diary and computer-
ized diaries, electronic diary and electronic
diaries, and ecological momentary as-
ssment. The search was limited to publica-
tions in English and the review is limited to
peer-reviewed journal articles.

In addition, other articles were located
by manual searches of bibliographies
from existing ED studies. This literature
search generated a total of 92 studies. For
the purpose of this review, we examined
only empirical studies published in the
peer-reviewed literature, which reduced
the number to 76 studies (Table 1). We
also identified 16 published papers dis-
cussing methodological issues in ED
research or other non-empirical discus-
sions of EDs, which were not included in
this review.

One important observation from the
empirical literature is how rarely
researchers report important details
about the implementation of EDs. For
example, one in four studies identified for
this review did not mention the manufac-
turer, type, or size of the ED. More impor-
tantly, half of the studies failed to report
compliance rates, and a similar proportion
provided no description of subject training
procedures. As others have noted, it is
important that ED researchers adopt
more thorough reporting standards.8

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Research area</th>
<th>N</th>
<th>Assessm'ts per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>Barr-Taylor et al.</td>
<td>CNS (Panic disorder)</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>1991</td>
<td>Rivelise et al.</td>
<td>Food consumption</td>
<td>21</td>
<td>—</td>
</tr>
<tr>
<td>1992</td>
<td>Haythornthwaite et al.</td>
<td>Respiratory</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>1992</td>
<td>Totterdell &amp; Folkard</td>
<td>Sleep</td>
<td>48</td>
<td>9</td>
</tr>
<tr>
<td>1993</td>
<td>Hylaned et al.</td>
<td>Respiratory</td>
<td>24</td>
<td>—</td>
</tr>
<tr>
<td>1993</td>
<td>Rabin et al.</td>
<td>Urinary incontinence</td>
<td>50</td>
<td>—</td>
</tr>
<tr>
<td>1993</td>
<td>Rosenflack &amp; Bendtson</td>
<td>Endocrinology</td>
<td>9</td>
<td>—</td>
</tr>
<tr>
<td>1993</td>
<td>Smith &amp; Safer</td>
<td>Pain (Chronic)</td>
<td>31</td>
<td>—</td>
</tr>
<tr>
<td>1993</td>
<td>Waldman et al.</td>
<td>Location tracking</td>
<td>6</td>
<td>—</td>
</tr>
<tr>
<td>1993</td>
<td>Zog et al.</td>
<td>CNS (Depression)</td>
<td>9</td>
<td>—</td>
</tr>
<tr>
<td>1994</td>
<td>Battig et al.</td>
<td>CNS (Addiction)</td>
<td>82</td>
<td>6</td>
</tr>
<tr>
<td>1994</td>
<td>Cox et al.</td>
<td>Endocrinology</td>
<td>41</td>
<td>—</td>
</tr>
<tr>
<td>1994</td>
<td>Hofer &amp; Battig</td>
<td>CNS (Addiction)</td>
<td>120</td>
<td>6</td>
</tr>
<tr>
<td>1994</td>
<td>Jacober et al.</td>
<td>CNS (Addiction)</td>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>1994</td>
<td>Lewis et al.</td>
<td>Pain (Arthritis)</td>
<td>36</td>
<td>4</td>
</tr>
<tr>
<td>1994</td>
<td>Nived et al.</td>
<td>Pain (Arthritis)</td>
<td>307</td>
<td>1</td>
</tr>
<tr>
<td>1994</td>
<td>Penner et al.</td>
<td>CNS (Mood)</td>
<td>54</td>
<td>—</td>
</tr>
<tr>
<td>1994</td>
<td>Shiffman et al.</td>
<td>CNS (Addiction)</td>
<td>57</td>
<td>5</td>
</tr>
<tr>
<td>1994</td>
<td>Stone et al.</td>
<td>CNS (Chronic Fatigue Syndrome)</td>
<td>29</td>
<td>5</td>
</tr>
<tr>
<td>1995</td>
<td>Clarke et al.</td>
<td>Endocrinology</td>
<td>78</td>
<td>—</td>
</tr>
<tr>
<td>1995</td>
<td>Drummond et al.</td>
<td>Gastrointestinal</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td>1995</td>
<td>Lewis et al.</td>
<td>Pain (Arthritis)</td>
<td>36</td>
<td>4</td>
</tr>
<tr>
<td>1995</td>
<td>Parkinson et al.</td>
<td>CNS (Mood)</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>1995</td>
<td>Shiffman et al.</td>
<td>CNS (Addiction)</td>
<td>51</td>
<td>5</td>
</tr>
<tr>
<td>1995</td>
<td>Totterdell et al. (a)</td>
<td>Menstruation</td>
<td>24</td>
<td>—</td>
</tr>
<tr>
<td>1995</td>
<td>Totterdell et al. (b)</td>
<td>Sleep</td>
<td>61</td>
<td>10</td>
</tr>
<tr>
<td>1996</td>
<td>Affleck</td>
<td>Pain (Fibromyalgia)</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>1996</td>
<td>Donovan et al.</td>
<td>Pain (Arthritis)</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>1996</td>
<td>Kos &amp; Battig</td>
<td>Food consumption</td>
<td>82</td>
<td>—</td>
</tr>
<tr>
<td>1996</td>
<td>Rabin et al. (a)</td>
<td>Urinary incontinence</td>
<td>50</td>
<td>—</td>
</tr>
<tr>
<td>1996</td>
<td>Rabin et al. (b)</td>
<td>Urinary incontinence</td>
<td>72</td>
<td>—</td>
</tr>
<tr>
<td>1996</td>
<td>Redelmeier &amp; Kahneman</td>
<td>Pain (Surgery)</td>
<td>287</td>
<td>—</td>
</tr>
<tr>
<td>1996</td>
<td>Shiffman et al. (a)</td>
<td>CNS (Addiction)</td>
<td>151</td>
<td>5</td>
</tr>
<tr>
<td>1996</td>
<td>Shiffman et al. (b)</td>
<td>CNS (Addiction)</td>
<td>133</td>
<td>5</td>
</tr>
<tr>
<td>1996</td>
<td>Shiffman et al. (c)</td>
<td>CNS (Addiction)</td>
<td>108</td>
<td>5</td>
</tr>
<tr>
<td>1996</td>
<td>VanCerven et al.</td>
<td>Pain (Migraine)</td>
<td>159</td>
<td>—</td>
</tr>
<tr>
<td>1997</td>
<td>Apter et al.</td>
<td>Respiratory</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>1997</td>
<td>King et al.</td>
<td>Food consumption/</td>
<td>8</td>
<td>15</td>
</tr>
</tbody>
</table>

The average ED study assesses sub-
jects’ experiences 7.2 (SD = 6.9) times per
day in the 56 (74%) studies that provide
details of their sampling. Once-daily
assessments, common for many paper
diary protocols, account for only 8% of
published ED studies. This is important

Table 1: Published studies of electronic patient diary use

In sum, the empirical literature sup-
ports the observation of many clinical
researchers who have been skeptical of
the veracity of paper diary data.8 As a
result of these historic misgivings about
paper diary data, combined with the need
for many trials to use diary data as pri-
mary and secondary endpoints,
researchers began to incorporate hand-
held technology to the collection of diary
data in clinical trials more than 10 years
ago.

**Literature review**

This literature review of EDs in clinical
research is based on a search of the
Medicine’s PubMed (Medline) databases
(www.ncbi.nlm.nih.gov/pubmed) and the
American Psychological Association’s
PsychInfo database (member only, but
abstracts can be viewed at www.apa.org/
psycinfo/). The search terms used for this
review were computer diary and computer
diaries, computerized diary and computer-
ized diaries, electronic diary and electronic
diaries, and ecological momentary as-
ssment. The search was limited to publica-
tions in English and the review is limited to
peer-reviewed journal articles.

In addition, other articles were located
by manual searches of bibliographies
from existing ED studies. This literature
search generated a total of 92 studies. For
the purpose of this review, we examined
only empirical studies published in the
peer-reviewed literature, which reduced
the number to 76 studies (Table 1). We
also identified 16 published papers dis-
cussing methodological issues in ED
research or other non-empirical discus-
sions of EDs, which were not included in
this review.

One important observation from the
empirical literature is how rarely
researchers report important details
about the implementation of EDs. For
example, one in four studies identified for
this review did not mention the manufac-
turer, type, or size of the ED. More impor-
tantly, half of the studies failed to report
compliance rates, and a similar proportion
provided no description of subject training
procedures. As others have noted, it is
important that ED researchers adopt
more thorough reporting standards.8

**EDs in clinical research**

The collection of studies in Table 1
reflects a broad array of applications. In
terms of the range of ED applications,
these studies range in duration from one
day to one year, with a mean length of ED
monitoring of 34.8 (SD = 65.9) days. The
mean age of subjects is 40.4, with a stan-
dard deviation of 8.5 in the 66 (87%) stud-
ies that report subjects’ ages.

The average ED study assesses sub-
jects’ experiences 7.2 (SD = 6.9) times per
day in the 56 (74%) studies that provide
details of their sampling. Once-daily
assessments, common for many paper
diary protocols, account for only 8% of
published ED studies. This is important
The hardware to support EDs in clinical research has become reliable, commercially available, and increasingly affordable. Many early studies relied on custom hardware for the ED. This was necessary at the time because few handheld computers were commercially available, and those on the market were not always highly reliable.

Over time, researchers have adopted standard handheld platforms for use as EDs. In the early 1990s it was common for researchers who have published the most ED studies to date, and subject preference for EDs.

As shown in Figure 1, the empirical literature reflects a clear increase over time in the application of EDs to clinical research. These empirical studies inevitably underestimate the true prevalence of EDs because many studies conducted in the pharmaceutical industry may not be published yet or may be currently underway.

Table 1 continued (for table references, see actmagazine.com).

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Research area</th>
<th>N</th>
<th>Assessments per day*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>Kos et al.</td>
<td>CNS (Addiction)</td>
<td>65</td>
<td>6</td>
</tr>
<tr>
<td>1997</td>
<td>Shiffman et al. (a)</td>
<td>CNS (Addiction)</td>
<td>214</td>
<td>6</td>
</tr>
<tr>
<td>1997</td>
<td>Shiffman et al. (b)</td>
<td>CNS (Addiction)</td>
<td>140</td>
<td>6</td>
</tr>
<tr>
<td>1997</td>
<td>Shiffman et al. (c)</td>
<td>CNS (Addiction)</td>
<td>105</td>
<td>5</td>
</tr>
<tr>
<td>1997</td>
<td>Shiffman et al. (d)</td>
<td>CNS (Addiction)</td>
<td>127</td>
<td>5</td>
</tr>
<tr>
<td>1997</td>
<td>Tiplady et al.</td>
<td>Respiratory</td>
<td>59</td>
<td>2</td>
</tr>
<tr>
<td>1998</td>
<td>Affleck et al.</td>
<td>Pain (Fibromyalgia)</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>1998</td>
<td>Carney et al.</td>
<td>CNS (Addiction)</td>
<td>48</td>
<td>—</td>
</tr>
<tr>
<td>1998</td>
<td>Collins et al.</td>
<td>CNS (Addiction)</td>
<td>37</td>
<td>7</td>
</tr>
<tr>
<td>1998</td>
<td>Finkelstein et al.</td>
<td>Respiratory</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>1998</td>
<td>Kamarck et al.</td>
<td>Cardiovascular</td>
<td>120</td>
<td>21</td>
</tr>
<tr>
<td>1998</td>
<td>O’Connell et al.</td>
<td>CNS (Addiction)</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>1998</td>
<td>Stone et al.</td>
<td>CNS (Addiction)</td>
<td>100</td>
<td>24</td>
</tr>
<tr>
<td>1999</td>
<td>Fajaj et al.</td>
<td>Gastrointestinal</td>
<td>94</td>
<td>—</td>
</tr>
<tr>
<td>1999</td>
<td>Honkoop et al.</td>
<td>Pain (Migraine)</td>
<td>56</td>
<td>6</td>
</tr>
<tr>
<td>1999</td>
<td>Kinne et al.</td>
<td>Cardiovascular</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>1999</td>
<td>Marco et al.</td>
<td>CNS (Coping)</td>
<td>100</td>
<td>24</td>
</tr>
<tr>
<td>1999</td>
<td>Milgrom et al.</td>
<td>Respiratory</td>
<td>283</td>
<td>2</td>
</tr>
<tr>
<td>1999</td>
<td>Schwartz et al.</td>
<td>CNS (Coping)</td>
<td>96</td>
<td>7</td>
</tr>
<tr>
<td>1999</td>
<td>Stephan et al.</td>
<td>CNS (Depression)</td>
<td>129</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>Affleck et al.</td>
<td>Pain (Fibromyalgia)</td>
<td>89</td>
<td>4</td>
</tr>
<tr>
<td>2000</td>
<td>Catley et al.</td>
<td>CNS (Addiction)</td>
<td>41</td>
<td>—</td>
</tr>
<tr>
<td>2000</td>
<td>Finkelstein et al.</td>
<td>Respiratory</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td>2000</td>
<td>Greeno et al.</td>
<td>CNS (Binge eating)</td>
<td>79</td>
<td>6</td>
</tr>
<tr>
<td>2000</td>
<td>Johannes et al. (a)</td>
<td>Menstruation</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>Johannes et al. (b)</td>
<td>Menstruation</td>
<td>25</td>
<td>—</td>
</tr>
<tr>
<td>2000</td>
<td>O’Connell et al.</td>
<td>CNS (Addiction)</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>2000</td>
<td>Peters et al.</td>
<td>Pain (Back)</td>
<td>80</td>
<td>6</td>
</tr>
<tr>
<td>2000</td>
<td>Porter et al.</td>
<td>CNS (Coping)</td>
<td>95</td>
<td>24</td>
</tr>
<tr>
<td>2000</td>
<td>Salmon et al.</td>
<td>Allergy</td>
<td>60</td>
<td>4</td>
</tr>
<tr>
<td>2000</td>
<td>Shiffman et al.</td>
<td>CNS (Addiction)</td>
<td>244</td>
<td>11</td>
</tr>
<tr>
<td>2000</td>
<td>Swendsen et al.</td>
<td>CNS (Addiction)</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>2000</td>
<td>Tiplady et al.</td>
<td>Respiratory</td>
<td>118</td>
<td>2</td>
</tr>
<tr>
<td>2001</td>
<td>Cox et al.</td>
<td>Endocrinology</td>
<td>73</td>
<td>2</td>
</tr>
<tr>
<td>2001</td>
<td>Jamison et al.</td>
<td>Pain (Back)</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>2001</td>
<td>Whalen et al.</td>
<td>CNS (Addiction)</td>
<td>153</td>
<td>30</td>
</tr>
<tr>
<td>In press</td>
<td>Hufford et al.</td>
<td>CNS (Addiction)</td>
<td>33</td>
<td>7</td>
</tr>
</tbody>
</table>

*Note that some of these studies represent multiple publications from the same dataset. A complete list of these references is provided in the Appendix.

The number of assessments completed per day was not reported in some studies. These values represent averages across subjects, and had to be estimated from summary data in the Methods sections of some of these publications.

In addition, the lag time associated with publication of papers also serves to delay the appearance of ED studies in the peer-reviewed literature. That EDs are being used more begs the question of what factors are facilitating their acceptance by researchers. Several factors are responsible for the increase in ED use over time.

**Hardware reliability.** The hardware to support EDs in clinical research has become reliable, commercially available, and increasingly affordable. Many early studies relied on custom hardware for the ED. This was necessary at the time because few handheld computers were commercially available, and those on the market were not always highly reliable.

Over time, researchers have adopted standard handheld platforms for use as EDs. In the early 1990s it was common for researchers to develop their own hardware for use as EDs in the field. In the past few years, this practice has become far less common.

**Software support.** The software required to support field use of EDs in subjects’ real-world environments has become increasingly stable. This is not just an inevitable consequence of conducting an ED study. The literature contains many examples of EDs whose software failed to perform as required (for example, uploading data regularly) or that allowed subjects to enter data retrospectively, negating the advantages of diary data in the first place.

Not surprisingly, researchers who have published the most studies—reflecting an increased maturity in the ED software—tend to have the highest rates of subject compliance. In short, experience counts in terms of the development of robust software for use in the field.

**Regulatory guidance.** Regulatory initiatives by the FDA, most notably the 1997 publication of 21 CFR 11 in addition to the creation of guidelines for electronic data collection systems, provided a framework to ensure that all electronic data can be collected to meet FDA data-integrity standards (defined by the principles that the data must be accurate, contemporaneous, attributable, and legible). As has been outlined elsewhere, an ED system must be

because it reflects the tendency of ED research to sample subjects’ experiences several times per day. Unlike paper diaries, which are often completed right before a site visit, EDs can be used to ensure timely completion of diary entries. Not all ED studies are equivalent in this regard. As outlined below, rates of subject compliance differ across studies.

Our survey of the ED literature touches on a number of topics, including the increased use of EDs in clinical research, the wide range of therapeutic areas that have published ED studies to date, and subject preference for EDs.

**Increasing use of EDs in clinical research.** As shown in Figure 1, the empirical literature reflects a clear increase over time in the application of EDs to clinical research. These empirical studies inevitably underestimate the true prevalence of EDs because many studies conducted in the pharmaceutical industry may not be published yet or may be currently underway.
developed and validated to FDA regulations (for example, 21 CFR 11) and also to
general industry standards so clinical trial
data can be submitted to the FDA. These
regulations have offered important guid-
ance and direction in the application of
EDs to clinical trials.

Subject compliance. As outlined in
detail below, the empirical data supports
the assertion that EDs enhanced with liv-
ability and compliance features can suc-
cede in driving high rates of subject com-
pliance to diary protocols. Subject
compliance is the key “vital sign” for how
well an ED system is performing. If the
user interface, hardware, software, or
another issue is causing problems for sub-
jects, the result will inevitably be a de-
crease in subject compliance.

Application across therapeutic categories. EDs have been used across a variety of
subject groups representing a broad range of therapeutic categories. Table 2
presents a breakdown of these studies by
category. The most active researchers
confirmed the number of studies per cate-
gory (Shiffman and colleagues, for exam-
ple, account for 27% of the published ED
studies listed in Table 1.)

In addition, EDs have been subject to
many primarily methodological inquiries
to establish their validity. Of the 76 studies
reviewed, 30% fall into this category. For
example, several studies have examined
reactivity to ED assessment. A reactive
effect is the degree to which the intensity,
frequency, and/or quality of a target vari-
able changes when being observed, moni-
tored, or assessed. Four studies have
examined whether real-time assessments
engender significant reactivity among
subjects. None of the four studies found
reactivity effects.

Other methodological studies have
examined burden of use and direct per-
formance comparison of EDs to paper
diaries. These studies found that EDs
produce higher quality, more reliable data
than their paper diary counterparts. Also,
ED data has been compared to data col-
lected at site visits to examine bias, inac-
curacy, and decreased sensitivity to medi-
cation effects of the retrospectively
collected data. For example, Nived
and colleagues showed that ED data could
detect in four weeks a medication effect
that took 24 weeks to emerge using data
collected episodically at site visits.

Subject preference for EDs. One question
that often arises regarding the application
of EDs to clinical research is whether sub-
jects prefer the more traditional paper
diaries to EDs. Several studies examined
subject preference for EDs and uni-
versally found that subjects readily accept
and even prefer these devices to paper
diaries. Drummond and colleagues
reported that 57% of their subjects with
gastrointestinal disorders preferred the
electronic assessment, and only 13% pre-
ferred the paper versions; 30% expressed
no preference. Furthermore, no associa-
tion with assessment preference was
found based on age, sex, or comfort with
technology and computer use. Tippley
and colleagues compared EDs and paper
diaries in 22 respiratory clinic outpatients
who self-monitored with both methods for
four weeks; 59% expressed preference for
the ED over the paper diary, 18% pre-
ferred paper, and 23% expressed no pre-
ference. Again, a subject’s age, sex, and
comfort or familiarity with technology
were not associated with their diary pre-
ference. Rabin and colleagues also
explored subject preference for EDs rela-
tive to paper diaries after allowing subjects
in a urinary incontinence study to self-
monitor with each method for one week.
Over 98% of their subjects and over 80% of
their control group expressed
an explicit preference for the ED over the
paper diary. Furthermore, both groups evaluated the ED more positively than the
paper diary on a variety of attributes (for
example, fun, easy to use, and feel
involved). Finally, Johannes and col-
leagues found that approximately 70% of
their all-female sample (n=23) preferred
an ED versus a paper menstrual diary.

Another study investigated whether
previous computer experience was a nec-
essary prerequisite to participate and be
compliant with EDs. Finkelstein and col-
leagues examined a sample of asthma
subjects from a low socioeconomic urban
community without previous computer
experience (n=17). Over three weeks, the
subjects were required to interact with the
ED at least twice daily. Results were that
the vast majority (82%) found the proce-
dures “not difficult at all.” This suggests
that previous computer experience is not
necessary for subject compliance. In sum,
the empirical literature reflects a clear

Figure 1. Graph of electronic diary publications by year (1990-2000).
subject preference for EDs relative to paper diaries.

Evaluating ED success

ED success can be evaluated in two primary ways. We use the term technical success to refer to whether the ED functioned as required in the field. More broadly, subject compliance reflects not only technical success (the ED must be functioning properly for data to be entered) but also whether the subject understood the ED training, how to use the ED, and when to provide data. In other words, technical success is a necessary but not sufficient prerequisite to subject compliance, who have developed a new ED solution for their research.

In terms of subject compliance, many researchers have noted that compliance data, despite its importance, is seldom reported in clinical trials—and often handled inaccurately. Standards do exist for the reporting of compliance in clinical trials, which consist of dividing the number of completed assessments by the number of scheduled assessments and multiplying by 100. When researchers do not follow these guidelines, the result can be an exaggeration of compliance rates with EDs. When evaluating compliance rates, it is important to avoid artifices, adequate subject training can certainly not be taken for granted. Finkelstein and colleagues asked their subjects if they felt they received all the necessary information about self-monitoring during the training session. Although most subjects reported receiving all the information necessary to complete the protocol, a significant minority of both samples (26% and 24%) indicated that they did not. The actual length of training was reported in only 9 of the 76 studies, and length of training ranged from less than one hour to three hours. No clear relationship emerged between subject compliance and the length of training.

Somewhat surprisingly, the number of assessments per day appears to be unrelated to subject compliance, even after controlling for the length of the monitoring protocol. For example, Kamarck and colleagues had 99% compliance with a protocol requiring 12 or more assessments each day. Their study used EDs with a number of compliance features and a short protocol to help balance out the intense subject burden during the monitoring period. Importantly, studies that sample subjects many times per day and have high rates of compliance demonstrate that researchers can use sophisticated EDs to do more than collect single assessments from subjects each day. More dynamic sampling protocols can be used to oversample critical periods, providing a sensitive test of medication effects as well as specific details of those effects, such as time-to-relief or onset-of-action hypotheses.

Some ED investigators have designed protocol components aimed specifically at increasing compliance, using them successfully in several ways and at varying levels of sophistication. For example, some researchers describe real-time feedback on the ED or regularly scheduled meetings throughout the course of protocols as excellent opportunities to provide subjects with detailed compliance reviews and feedback and to create a sense of accountability for the data they are providing.

The future of electronic diaries

The future of EDs in clinical research surely holds many promises and challenges. In terms of the promise of EDs, studies published to date clearly show that
The term for these methods, ecological momentary assessment (EMA), has been widely applied in academic research and more recently used in pharmaceutical clinical trials.

Certainly, ED technology will continue to evolve at a rapid pace, offering improvements in computational power, screen resolution, on-board communication, and built-in features—and producing a wide array of impressive handheld hardware from which to choose. Another exciting area of innovation is the integration of ambulatory monitoring with EDs. But this review of EDs in clinical research clearly suggests that fancy hardware alone is not enough to ensure success in the field. Robust hardware is a necessary, but not sufficient, precondition to a successful ED. In other words, subject compliance with EDs shows that technology alone is insufficient to ensure successful data capture in the field. Researchers need to differentiate between technological novelty and true innovations that allow better data to be collected from subjects in a clinical trial. In short, the technology that shows subjects’ preference for EDs, and empirical data that verifies high rates of subject compliance in many ED protocols. The peer-reviewed literature also shows that EDs are being used to sample subjects’ experiences many times per day, providing researchers with new insights into the effects of experimental medications.

Clearly, not all ED studies are a success in the field. The variation in terms of both technical success and subject compliance appears to be a function of the maturity of the ED system, the robustness of the subject training, and incorporation of scientifically based compliance features that encourage subjects to use the devices in their daily lives.

References
1. L.M. Verbrugge, “Health Diaries,” Medical Care, 18, 73–95 (1980).

References from Table 1


Michael R. Hufford, *PhD, is director of Scientific Affairs at invivodata, inc., 2100 Wharton St., Suite 505, Pittsburgh, PA 15203, (412) 390.3008, fax (412) 390.3020, email: mhufford@invivodata.com. Alan L. Shields, MA, is at the Department of Psychology, University of Montana.

* To whom correspondence should be addressed.