

Research Article

Reducing Adverse Drug Reactions for Older People in the Community: Evaluating the Validity and Reliability of the ADRe Profile

Vera Logan ¹, Neil Carter ², David Hughes ², Adam Turner ² and Sue Jordan ²

¹Faculty of Health and Social Sciences, University of South Bohemia, Ceske Budejovice, Czech Republic

²Faculty of Medicine, Health and Life Science, Swansea University, Swansea, UK

Correspondence should be addressed to Vera Logan; vlogan@zsf.jcu.cz

Received 31 December 2024; Accepted 12 April 2025

Academic Editor: Abdulqadir J. Nashwan

Copyright © 2025 Vera Logan et al. Journal of Nursing Management published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Background: Adverse drug reactions (ADRs), particularly in the context of polypharmacy, remain a persistent, unresolved problem for patients and healthcare professionals. The ADRe Profile identifies medicine-related harms, and supports their resolution, thereby improving care quality and preventing future problems.

Objective: The objective of this study was to assess the validity and reliability of the ADRe Profile (<https://www.swansea.ac.uk/adre/>) in U.K. primary care general practices, building on assessments in other settings.

Methods: The ADRe Profile's validity and reliability were investigated using complementary mixed methods: content validity index, contrast group construct validity, cognitive interviewing, and inter-rater reliability.

Results: Cognitive interviews ($n = 5$) confirmed that the ADRe Profile needed only minor adjustments. The scale-level content validity index was 0.67 ($n = 14$), items ranging from 0.08 to 1. Significant differences in signs and symptoms associated with ADRs between service users taking different numbers of regular prescribed medicines confirmed construct validity ($n = 68$, $U = 870.50$, $p < 0.001$). Inter-rater reliability testing showed substantial agreement between service users and research nurse: 10 items had 100% agreement. Overall κ mean was 0.71 (range: 0.31–1), ($n = 42$).

Conclusions and Relevance: The ADRe Profile is suitable for use with older service users in primary care who live at home. Users understood the questions and provided meaningful answers. ADRe Profile responses were sufficiently reliable to be used as a basis for further investigations, prescriber referral and clinical actions. However, clinician judgement of content validity may depend on knowledge and experience, highlighting the importance of training. Clinicians acknowledged that the ADRe Profile was comprehensive but identified practical difficulties. Instruments to reduce ADRs should be validated before testing in feasibility studies and randomised controlled trials.

Implications for Nursing Management: Managers need to optimise patient safety by introducing patient-centred symptom monitoring, with decision support. Before instruments are adopted, managers should check the reliability and validity data.

Trial Registration: ClinicalTrials.gov identifier: NCT04663360

Keywords: drug-related side effects and adverse reactions; pharmacovigilance; primary health care; validation study

1. Introduction and Background

Adverse drug reactions (ADRs) are defined as ‘appreciably harmful or unpleasant reactions, resulting from an intervention related to the use of a medicinal product, which

predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product’ [1], p. 1255. The ADRe Profile (<https://www.swansea.ac.uk/adre/>) is an interdisciplinary instrument based on an ADR checklist

complemented by decision support, encompassing information on potential aetiology, additional assessments and strategies for monitoring or addressing problems at the click of an icon.

As more people are prescribed medicines for long-term conditions, the problem of ADRs is increasing [2, 3]. ADRs are pervasive: A systematic review and meta-analysis in primary care revealed an ADR prevalence of 28.43% among the elderly (> 65 years) [4]. In the United Kingdom, ADR-related unplanned hospital admissions count for 8.3% of the total [5]. ADR-related admissions and bed-days increased by over 50% between 2008 and 2015 [6]. Systematic reviews estimate ADRs affect one in 6 hospitalised patients [7], rising to 85% among elderly individuals with dementia [8]. ADRs attributable to error cause 712 deaths directly and contribute to another 1708 each year [9]. These statistics have generated interest in the monitoring and management of ADRs.

Current approaches to monitoring ADRs or side effects of medicines are unable to mitigate the harm emanating from ADRs [10, 11]. Medicine chart reviews [12, 13] and patient self-reports are the most common approaches to monitoring ADRs; although these improve prescribing, there is little impact on patient outcomes [14–16]. There is an urgent need to develop better ways of identifying and reducing ADRs to control this hidden epidemic [2]. However, to ensure that volunteers' time contributes to worthwhile developments, relevant instruments must be validated with their target population before implementation in clinical trials and practice [17].

Reliability and validity are pivotal to assessment of instruments that measure changes in health status or quality of life [18, 19]. It is difficult to claim clinical effectiveness for healthcare interventions if information on their reliability and validity is not in the public domain [20, 21]. The evaluation of measurement techniques allows a comprehensive understanding of the potential biases inherent in the utilization of specific instruments [22]. Absence of validity data risks meaningless and inaccurate outcomes [23]. Inferiorly executed randomised controlled trials (RCTs) are more likely to yield biased outcomes than well-conducted RCTs or nonrandomised studies [24]. Accordingly, the development of the ADRe Profile has followed a trajectory of iterative refinement and publications relating to reliability and validity [25–34]. Initially, the checklist items were assembled from evidence-based sources, mainly the British National Formulary and manufacturers' summaries of product characteristics (SmPCs), and curated by a multidisciplinary team comprising expert doctors, pharmacists, academics and nurses. Subsequent iterations involved ongoing adjustments to both the items and the accompanying supporting information, informed by implementation experience.

Before using the ADRe Profile in a RCT evaluating the clinical impact of its implementation in older polymedicated adults living in their own homes, validity, reliability and feasibility needed to be established in this group. study protocol is published [35].

This paper describes evaluation of the validity and reliability of the ADRe Profile. Feasibility testing, undertaken

to ascertain whether the ADRe Profile could be implemented in practice, and the trial results will be reported in future publications.

2. Methods

2.1. The Study Design. To evaluate the reliability and validity of the ADRe Profile, both qualitative evidence and quantitative evidence were gathered. Content validity was established through cognitive interviews and the calculation of a content validity index (CVI). Construct validity, which assesses whether the concept's application relates to the number of medicines taken, rather than other situational factors, was confirmed by implementing the ADRe Profile in contrast groups. Inter-rater reliability (IRR) was measured with Cohen's kappa coefficient using data from both the feasibility test and the intervention arm of the RCT. The evaluation of the ADRe Profile's validity and reliability was conducted in accordance with the definitions provided by the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) [36] and the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) [37]. A flow chart detailing the sequence of tests conducted is provided (Figure 1).

2.2. Setting and Participants. Participants were recruited in Southwest Wales between April 2021 and July 2022, through advertising in the Wales Centre for Primary and Emergency Care Research (PRIME), research networks and Swansea University. Cognitive interviews were undertaken with service users aged 65 or over and healthcare professionals in primary care who might use the ADRe Profile. A CVI was then calculated. Construct validity testing involved two groups of 20 and 48 service users aged 65 or more: one group without any regular medicines and a second group prescribed five or more items daily. The latter were recruited from six general practices in Southwest Wales; these participants completed the ADRe Profile with one researcher (VL). IRR testing was calculated with 42 patients using 5 or more regular medicines from 5 practices (see Figure 1 for an outline). These participants were recruited from the feasibility study plus a further three general practices in Southwest Wales. (Further details are described in [35]).

2.3. Ethical Approval. Ethical approval was granted by the Health Research Authority Research Ethics Committee on 17 March 2021 (Integrated Research Application System [IRAS] 292693). All participants gave informed, signed consent to their participation.

2.4. Cognitive Interviews. To ensure the accuracy and functionality of the ADRe Profile, cognitive interviews were conducted following established procedures [38–40], in accordance with the methodological framework of the Cognitive Interviewing Reporting Framework of Boeije and Willis [41]. Think-aloud/read aloud techniques, augmented

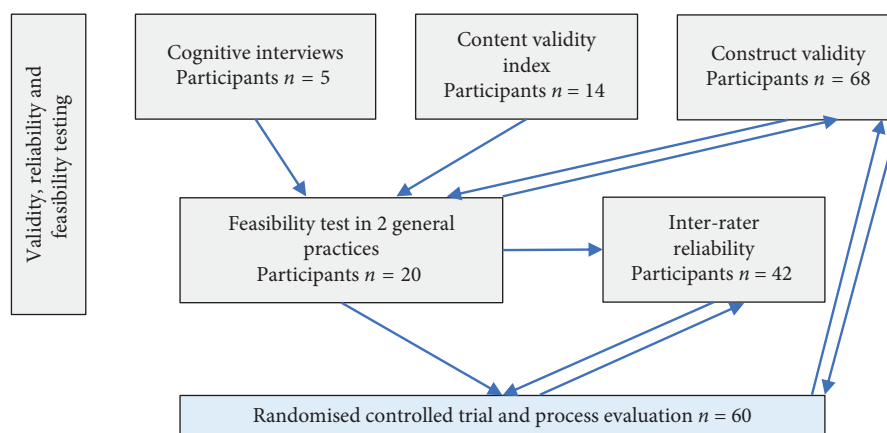


FIGURE 1: Flowchart of reliability and validity testing.

by concurrent probes [38, 42], were employed, and interpretations of the questions were noted.

Five participants were approached for cognitive interviews, and all agreed. The determination of sample size considered:

1. The nature of the questionnaire, whereby concrete questions about the respondent's health and well-being may necessitate fewer cognitive interviews than exploration of more abstract issues [43].
2. Prior cognitive interview testing of earlier versions of the ADRe Profile, with almost the same content [31, 32].

The sample was matched to the intended target population of users and recipients of the ADRe Profile. Two patients aged > 64, a nurse, a pharmacist and a general practitioner (GP) participated. Interviews were conducted by a single researcher (VL) via the Zoom video conferencing platform at mutually convenient times between April and October 2021. Face-to-face interviews were not possible due to the pandemic.

The interview protocol underwent refinement following a preliminary 'test run' within the research team [44]. An audio recording of each interview was made to facilitate thorough analysis, while the interviewer maintained detailed notes to capture immediate insights. Since Willis [45] suggests that cognitive interviews are not typically transcribed, we relied on detailed notes and recordings to capture participants' responses.

Comments pertaining to specific questionnaire items were coded into predefined categories, informed by the Question Appraisal System (QAS-99) [46]; any responses not aligning with the predetermined categories were labelled as 'other', and thematic reduction techniques were applied to identify overarching themes. Data reduction techniques [47] were employed to condense and organise the data into meaningful categories.

2.5. CVI. To ascertain whether the items on the ADRe Profile represent health and well-being problems associated with multiple medicines commonly prescribed for older individuals receiving primary care, a CVI was calculated,

following established methods [48–50]. From professional contacts within the Faculty of Medicine, Health and Life Science, a patient volunteer group and professional contacts for community practitioners, we recruited nurses working in general practice, GPs, pharmacists, and care assistants working in general practice/community and older patients. Data were collected through an electronic survey distributed upon completion of the consent form. Participants were instructed to rate the ADRe Profile items on a 4-point ordinal scale [51], as follows:

1. Not relevant
2. Somewhat relevant
3. Quite relevant
4. Highly relevant

The 101 items on the ADRe Profile could score a maximum of 404 (if all items were rated as highly relevant) and a minimum score of 101 (if all items not relevant). Participants were also given an opportunity to comment on any items they deemed necessary for removal or addition, as well as to add any other comments in free-text format.

The CVI was calculated for each questionnaire item and represented the strength of agreement among participants regarding the relevance of the item (i.e., scoring 3 or 4 on the scale), as per Lynn [51]; Waltz and Bausell [52]; and Polit and Beck [53]. The multirater kappa statistic (κ) was calculated to determine the level of agreement beyond chance [54]. Kappa values were interpreted based on Landis and Koch's [55] criteria: > 0.80 indicating almost perfect agreement, 0.61–0.80 substantial agreement, 0.41–0.60 moderate agreement, 0.21–0.40 slight agreement and < 0 poor agreement, below the level of chance. The scale content validity ratio was calculated as a proportion of total items judged against content valid items and as the final average of individual CVI scores (as in [49, 56]). Free-text comments were analysed qualitatively.

2.6. Contrast Group Construct Validity. Construct validity was investigated through the comparison of contrast groups, with the expectation that their respective outcome measures

would have different scores on a measurement scale [57, 58], taken here as the number of issues identified on the ADRe Profile. The contrast groups were constructed based on the quantity of prescribed medicines (no medicines vs. 5 or more medicines daily).

A total of 68 participants were recruited, 20 in the 'no-medicines group' (65 years old or older, using no prescribed medicines (vitamins, nutritional supplements and moisturising skin preparations were not counted as medicines) and 48 in the 'polypharmacy group' (65 years old or older, 5 or more long-term medicines prescribed by GPs). Information on the number of daily prescribed medicines for the 'polypharmacy group' was extracted from the medical notes [35].

Patients were approached by clinicians and consented by researchers between May and December 2021.

All participants completed the ADRe Profile, either independently (via online platform or on paper) or with the researcher over the telephone. The sections on vital signs and medicines administration were omitted, leaving 57 items in the assessment. The Mann–Whitney statistic was used to test differences between the groups.

2.7. IRR. IRR was calculated based on self-administration followed by researcher-led administration of the ADRe Profile in person or via video link with 42 patients aged 65 or over prescribed at least 5 long-term medicines. Self-completion preceded researcher administration by 1–7 weeks.

The IRR of patient- versus researcher-completed ADRe Profiles was measured to ascertain the extent to which potential variability in records was attributable to random measurement errors [59, 60] and estimate the degree of agreement between patient and researcher evaluation of patients' symptoms. Additionally, the sensitivity and specificity of the ADRe Profile were calculated as (true positives)/(true positives + false negatives) and (true negatives)/(true negatives + false positives) [61, 62].

A minimum sample size of 50 participants was deemed sufficient for scenarios involving two raters and binary endpoints, to allow 2-sided testing against a null value of 0.00 for kappa with 80% power and alpha of 0.05 [63, 64].

For categorical ratings, the proportion of agreement between the two raters was calculated using the following equation: (number of agreements)/(number of agreements + disagreements) [65]. Cohen's kappa (κ) was calculated as described in [66], using the equation below:

$$\kappa = \frac{\text{proportion of agreement} - \text{probability of chance agreement}}{1 - \text{probability of chance agreement}} \quad (1)$$

All participants received guidance on using the ADRe Profile, to improve rating consistency. To further minimise systematic error, brief written completion instructions and structured response formats with predefined answers were included.

3. Results

3.1. Cognitive Interviews. Analysis of the cognitive interview comments and suggestions led to 15 adjustments to the ADRe Profile. Eleven adjustments were made to the 'Supporting information'; for example, 'muscle cramps' [67] and 'crossing and uncrossing legs when sitting' [68] were added to the list of manifestations of 'Abnormal movements at rest'. Three adjustments were made in the ADRe Profile. For example, 'More than 2 doses of prescribed medication missed over any period of 7 days in the last month' was replaced with 'More than 2 doses of prescribed medication missed in the last 7 days', to increase clarity. One adjustment was made in the 'How to' document. Detailed information regarding these adjustments is available in Supporting Information 1.

The cognitive interviews gave participants the opportunity to reflect on the ADRe Profile. All participants acknowledged the usefulness and need for such a tool. However, the healthcare professionals raised concerns regarding the time needed to complete the instrument and

address any problems identified. In addition, the GP expressed apprehension regarding medicolegal vulnerability:

GP: 'But that also slightly fills me with a little bit of fear. If I end up with this questionnaire on my table that says the patient has got chest pain, shortness of breath, vertigo, skin problems, erectile problems, and mental health problems, and behavioural changes. I fear having all that information on the table, and now I am responsible for actioning this, all this information. You can't . . . un-know that information, you've got it there, documented in writing. You put it into the patient's notes, because we are told we have to record everything and suddenly it looks like we are ignoring (smiles apologetically) all these kinds of concerns that patients have got.'

Other respondents were pleased to see potentially sensitive questions presented for discussion:

Pharmacist: *It is useful to include questions such as self-harm, physical aggression, violence, libido, and sex drive in the questionnaire, because such questions are rarely asked during medicine reviews. They are sensitive questions, which could be seen as judgemental if asked, but when included in a questionnaire, they can be seen as justified and it is easier to ask them. I don't normally ask about those*

things, and I bet you in your standard consultations no-one else would, so maybe there is actually an opportunity that someone would finally ask a question about that.

The practicalities of introducing a holistic, comprehensive approach to monitoring polypharmacy were raised:

GP: *'We certainly don't have an hour or half an hour or 20 minutes to do that. We should do, especially for our patients with polypharmacy and multiple, multiple problems but we just don't have that luxury, just because of the demand on our service, basically. So unfortunately, it should be a comprehensive chat about the medication, but it tends to be a quick 'how are you doing on your medication, are you taking it? Yes/no. Any problems? No, OKay, well, you know, can you reduce your PPI [proton pump inhibitor] because it is causing you calcium absorption problems or whatever. OK, I'll try, thank you very much, bye bye.'*

Pharmacist: *'Specialist reviews (diabetic, asthma, blood thinning medicines) take place in the GP practice. Each of the reviews only deals with problems pertaining to a single speciality: if someone has, for example, a respiratory condition, diabetes and has 10 other tablets for different conditions, I don't touch the respiratory [medicines], I don't touch the diabetic [medicines SIC], because we have our respiratory nurse, our diabetic nurse, and it's about 20 minutes each to do those reviews properly'.*

This raises questions about the fragmentation of care and implies difficulties in finding the space/time for implementing the ADRe Profile.

A comprehensive overview of the cognitive interview feedback is included in Supporting Information 2.

3.2. CVI. Recruitment was by staff email to ~150 academics in the Faculty of Medicine, Health and Life Science and other professional contacts. Fourteen participants were recruited to the CVI testing exercise: 5 GPs, 3 nurses, 2 pharmacists, 2 support workers and 2 patients. All participants completed the CVI survey online.

Among the 101 items in the ADRe Profile, two items ('dizziness' and 'falls') achieved unanimous agreement (14/14); 37 items met the generally accepted validity threshold, with a CVI of 0.78 or higher. The research team closely examined the 64 items with lower CVI values, leading to the removal of two items with CVI < 0.3 ('girth measurement' and 'are your drinks sugar-free?'). The calculated kappa values closely mirrored the CVI calculations, illustrating the diminishing effect of chance agreement with an increasing number of raters and the subsequent convergence of CVI and kappa values [69]. The scale-level index of the ADRe Profile was 0.67, with 37.4% of items demonstrating a significant level of agreement. Additionally, 85% of items had a CVI equal to or higher than 0.5. Table 1 provides a summary of the results; full results can be found in Supporting Information 3.

The 'Medicines administration' section of the ADRe Profile had the largest proportion of items with I-CVI > 0.78 (4/6), followed by the 'Vital signs' section (8/18), 'Observations' section (7/18), 'Reports/questions' section (16/42), 'Summary' (1/3) and 'Health promotion' section (4/15). The calculated kappa values were very similar to the I-CVI calculations and demonstrated the diminishing effect of chance agreement with increasing number of raters and subsequent converging of I-CVI and kappa [69].

The traditional vital signs were rated highly, but, despite medicines causing weight changes and firm guidance in the British National Formulary [70], the item of 'weight change' was not considered important. Only some posture and movement signs were rated highly, despite clinicians' reliance on pattern recognition. Pain, patients' views and health promotion were judged unimportant.

Participants' comments provided insight into their understanding of the ADRe Profile items. The mean time spent completing the CVI survey was 12 min (SD = 11.73). Given the instrument's complexity, some items may not have been fully understood. For example, CVI participant 2523 (a health care support worker) stated:

'The section on medicines management, don't [sic] make much sense to me. Looking at the first question '(are) any tablets crushed or broken?' What exactly are you asking here?'

This has likely influenced the CVI results. Another respondent's comment confirmed the need for the ADRe Profile and patient safety training:

'we normally do not do such a structured and detailed review I would say, so that means that people haven't been previously trained on the ADRe Profile documentation' [a pharmacist].

3.3. Contrast Group Construct Validity. Sixty eight participants contributed data. The 'no-medicines group' and the 'polypharmacy group' participants had comparable demographic characteristics: Median (IQR) ages were 68.50 (65.25–74.00) and 73.50 (69.00–78.00); 45% and 44% were female. Group differences in the numbers of problems identified on the ADRe Profile are summarised in Supporting Information 4. There was a significant difference in the median number of problems stated ($U = 870.50$, $p < 0.001$).

Certain ADRe Profile items, such as 'tongue tremor' and 'self-harm', were not reported by any participants in either group. Most items ($n = 27$) were not reported by participants in the 'no-medicines group' but were present in the 'polypharmacy group'. Examples include 'feet shuffling' ($n = 11$ [23%]) and 'problems with balance' ($n = 18$ [38%]). Several issues were encountered by participants in both groups, but the proportion of participants experiencing these issues in the 'no-medicines group' (n_1) was lower than in the 'polypharmacy group' (n_2). Examples include 'pain' ($n_1 = 7$

TABLE 1: Overview of the ADRe Profile content validity index results.

ADRe section	Section S-CVI	Modified kappa—section mean
Vital signs	0.62	0.62
Observations	0.69	0.69
Reports, questions	0.72	0.72
Prevention and health promotion	0.50	0.50
Medicines administration	0.83	0.83

[35%] vs. $n_2 = 37$ [77%]), indigestion ($n_1 = 2$ [10%] vs. $n_2 = 20$ [42%]) and 'sleep problems' ($n_1 = 5$ [25%] vs. $n_2 = 15$ [31%]).

3.4. IRR. Most (42 of 48) eligible participants (from 5 GP practices) returned their self-completed ADRe Profiles. One of the six noncompleters struggled to read even large print due to advanced cataracts, two disliked completing paper forms, and three participants postponed/neglected form completion. IRR was calculated for 78 questions presented in both the self-administered and researcher-administered ADRe Profile. Vital signs and the patient summary were not included in the IRR testing. Agreement was complete for factual statements, for example visits to dentists, and lower for less precisely defined questions, for example skin or respiratory problems. And, 10 of 78 items had complete agreement. The findings are summarised in Table 2, and kappa results per category can be found in Supporting Information 5.

The simple percentage agreement ranged from 69% to 100%, with a mean of 91.45% (SD = 6.92) and a median of 93 (25th percentile: 86%, 75th percentile: 97%). The aggregate kappa values indicated substantial agreement beyond chance. The full IRR report by item may be found in the Supporting Information 6. Given the recognised effects of prevalence bias and natural fluctuations of signs and symptoms over 1–7 weeks [71], these values were considered satisfactory and denote significant concordance in the service user and researcher ratings of items in the ADRe Profile. Sensitivity and specificity, unaffected by problem prevalence, were notably high (with mean sensitivity and specificity for the researcher rating at 85% and 90%, respectively).

4. Discussion

Evaluation of instrument validity and reliability contributed to development of the ADRe Profile for primary care use (see also: [72]). Validity and reliability were supported through complementary mixed measures: Construct validity was indicated by the CVI of 0.67 and cognitive interviews; discriminant or contrast validity indicated a difference in reports between service users not prescribed regular medicines and those prescribed five or more ($U = 870.50$, $p < 0.001$); IRR was demonstrated for all items (mean kappa 0.71, range [0.31–1.00]).

4.1. ADRe Profile Validity and Reliability. Content validity was assessed by confirming that the ADRe Profile items reflected the notion of ADRs and suboptimal medicines management, using expert ratings (CVIs) and cognitive

interviews to establish the CVI [51, 69]. Construct validity was evaluated by implementing the contrast group approach, comparing individuals taking no medicines versus those taking five or more, to assess whether the ADRe Profile could effectively distinguish between groups with expected differences in ADR-related health issues [57]. This use of 'known groups', where only one group could possibly have ADRs, offers an external measure of discriminant or criterion validity [73].

The assessment of validity and reliability properties was approached within a continuous framework [59] rather than as dichotomous outcomes. This approach implies that the accumulation of evidence increases confidence in the instrument's quality and the value of the data it generates.

The cognitive interview testing confirmed that most of the ADRe Profile items were clearly understood by the target population and appropriate for their situation. Testing of previous versions of the Profile [32] resulted in gradually diminishing numbers of implementable suggestions, indicative of progressive refinement of the instrument.

Favourable evaluations with cognitive interviews were corroborated by findings from the construct validity evaluation. Consistent with existing literature [57, 58], administering the instrument to contrast groups elicited divergent scores. The ADRe Profile identified significantly more problems among people using 5 or more medicines compared with those who had no regularly prescribed medicines.

Regarding content validity, just over a third of the questionnaire items met the generally accepted validity threshold (CVI of 0.78 or above), contrasting with previous findings of high scale CVI of 0.99 [32], or 93 out of 102 items rated as of the highest importance to delivery or outcomes of care [31]. It is unlikely that small differences in ADRe Profile versions account for this discrepancy, suggesting potential variations in raters' perspectives. While explanation for all items is included in the Supporting information part of the ADRe Profile, busy raters may not have time to read this resource sent to them in the ~12 min they allotted to completing the CVI; therefore, ratings may not have been fully informed.

Previous ADRe Profile effectiveness studies, which consistently demonstrated clinical benefit [26, 29, 31, 33, 34], and literature and clinical guidelines support the inclusion of all ADRe Profile items. User judgement of content validity may depend on knowledge and experience. Knowledge of pharmacology and, possibly, ADRe Profile training, may be crucial in future use.

The ADRe Profile demonstrated consistency across patient self-report and nurse researcher completion, as evidenced by high agreement and substantial aggregate

TABLE 2: Kappa frequency statistics, overall and per section.

ADRe profile items (<i>n</i> = 78)	Valid	Missing	Mean kappa value	Std. deviation	Range (min–max)
Overall	70	8	0.71	0.17	0.69 (0.31–1.00)
Observation section	17	1	0.72	0.16	0.62 (0.31–0.93)
Reports section	38	3	0.73	0.16	0.62 (0.38–1.00)
Health promotion section	12	1	0.69	0.22	0.68 (0.32–1.00)
Medicines section	3	3	0.60	0.05	0.10 (0.55–0.66)

kappa values during IRR testing, and high sensitivity and specificity results, comparable to previous testing findings [27, 32].

Although some ADRe Profile items are measurable (e.g., blood pressure or weight) or observable (e.g., gait or swelling), others rely on service users' reports of symptoms (e.g., pain, tiredness or forgetfulness). This poses challenges associated with the reliability and stability of self-reported data [74–77]. While literature specifically examining self-reporting health issues yields variable results [78–80], the IRR determined that participant ADRe Profile responses are sufficiently stable for meaningful use in clinical practice. Reliability testing empowered patients to recognise some clinical problems that had developed insidiously over some years and empowered them to seek help. For example, patients were helped by recognition that tremors plus abnormal movements may be pathognomonic of drug-induced Parkinsonism in patients using gabapentinoids, or that bleeding may indicate a need to reduce doses of anticoagulants.

4.2. The Need for High-quality Instruments Identifying ADRs. Questionnaires designed to identify ADRs should include a combination of closed and open questions, providing opportunities for service users to add comments and requests. Relying solely on open-ended questions such as: 'Do you experience any side effects?' assumes that patients have a comprehensive understanding of ADRs' symptoms. Therefore, questionnaires must be comprehensive enough to detect all or most ADR signs and symptoms. Since many commonly prescribed medicines each have many potential ADRs, for example there are ~100 'undesirable effects' for risperidone [81] and ~80 for gabapentin [82], it is difficult to reduce the list of potential problems whilst retaining a focus on safety. Healthcare professional respondents indicated that while the ADRe Profile is comprehensive, it is also perceived as too long and identifies clinical problems that professionals need to find time to manage. This may pose challenges for completion within the constraints of U.K. NHS practice. However, this downside is likely to be balanced by the clinical gains stemming from thorough ADR identification and the early resolution of issues [27, 83].

The WHO ([84], p. 8) notes that current guidelines focus on single diseases, reducing their effectiveness. As a pharmacist respondent (P1) noted, the ADRe Profile covers a wider range of conditions, and must consequently take the form of a longer instrument requiring significant completion time. The ADRe Profile includes problems potentially

related to a range of medicines and offers decision support relating to commonly prescribed medicines. It meets the WHO criteria for management of polypharmacy in that it implements medication reviews in collaboration with the patient ([85], p. 10). However, as our interviewees noted, close engagement with patients poses considerable difficulties given the current demand and resource pressures affecting U.K. primary care, and these will be difficult to overcome in the short term, without managerial commitment to reduce the number of unplanned hospital admissions due to ADRs.

4.3. Costs. Decision aids, such as the ADRe Profile, enhance professionals' perceptions of knowledge and judgement, but extend consultation time, particularly on first use [86, 87]. Costs of this extended patient contact are estimated as £14 for 15 min of a pharmacist's time, plus £4 for 10 min for a support worker to record vital signs; where an additional, unscheduled GP review is needed, this costs £32 [88]. Many of the problems addressed by ADRe Profile administration (e.g., relief of pain, dyspnoea or sedation) benefit patients but are difficult to cost. Many patients had several problems that could lead to hospitalisation. The quantifiable economic benefits of the ADRe Profile may be related to identification of these problems and prevention of admissions [29]. For example, each admission due to a fall, without a prolonged stay, costs ~£18,000 [89]. Therefore, the costs of the ADRe Profile would be recouped were it to prevent 1 falls' admission for each 250–730 administrations. Simultaneously, if pre-emptive action prevents bleeding escalating to an upper gastrointestinal haemorrhage, costing ~£3000 [90], the costs of 40–120 ADRe Profile administrations would be recouped.

4.4. The 'ADR Problem'. The 'ADR problem' continues to cost lives, cause preventable suffering and strain health services' resources [84] but has not been solved by existing strategies, such as medication reviews [10, 11, 14–16, 31]. Reactive, post hoc, voluntary reporting of serious ADRs (such as the U.K. Yellow Card scheme), whilst essential for postmarketing pharmacovigilance, does not benefit affected patients. Reviewers estimate voluntary reporting overlooks 94% of events [91], whilst in practice, < 1% of drug-induced GI haemorrhages are reported in the United Kingdom [92]. The role of pharmaceutical manufacturers in analysing [93] and disseminating ADR information may be suboptimal [94]. Accordingly, comprehensive instruments need to be incorporated into routine practice. Patient-reported

outcome questionnaires, with decision support, are a useful way of gathering comprehensive ADR data to complement medicine reviews.

Currently, most patient-reported questionnaires focus on specific medication classes, mostly mental health medicines, and lack decision support, for example the Glasgow Antipsychotic Side-Effect Scale [95], or the Maudsley Side-Effects Scale [96]. Few instruments target commonly prescribed medicines and, where they do, they may not be comprehensive [97]. Moreover, many patient-reported questionnaires lack critical quality indicators: A systematic review identified 19 patient-reported outcome instruments, but only half had been assessed for validity (50%) and even fewer for reliability (< 50%), with very few tested in trials [98]. Before clinical implementation, instruments must be supported by publicly funded evidence of reliability, validity, feasibility, and clinical effectiveness: A template is offered in this paper.

5. Strengths and Limitations of This Study

This work builds on earlier reports of reliability and validity of the ADRe Profile [27–33]. Combining the methods of testing offers an overview of instrument testing and leaves an overall impression as to the value of the ADRe Profile. The feasibility of administering the ADRe Profile is reported in tandem with this paper.

The ADRe Profile does not generate a scale or a score; rather, it lists individual problems. To our knowledge, there are no ‘gold standard’ comprehensive ADR scales for polypharmacy to which the ADRe Profile might be compared to test criterion validity [73]. It was not possible to verify all participants’ reports of symptoms to external testing. Traditionally, clinicians rely on patient self-report of symptoms [99]. Earlier ADRe studies have shown how acting on these self-reports reduced symptoms, such as pain or aggression. Reports of vision difficulties were externally validated when opticians prescribed spectacles; similarly, pain symptoms were externally validated when dentists extracted or filled teeth [28, 29].

The relatively low numbers of participants for some tests reflects the iterative nature of ADRe Profile refinement and available resources. However, similar findings to previous work indicate the practical adequacy of the ADRe Profile [100].

The validity and reliability testing had both strengths and limitations. The cognitive interviews adhered to recognised frameworks in planning, conducting, analysing and reporting [41, 47, 101]. However, the sample size was small and the conventional qualitative indicator for data collection cessation—data saturation—was not pursued. It is, therefore, possible that further interviews would collect data not captured in this study.

The construct validity testing in this study is specified relative to the construct that people who take 5 or more regular medicines exhibit more ADRs than people who do not. However, the aetiology of health and wellbeing problems is often multifactorial, making it challenging to distinguish between issues stemming from ADRs, undertreated health conditions, or symptoms of physiological ageing.

The limitations of IRR testing were mostly attributable to imperfect adherence to some of the scenario assumptions in Cohen’s kappa determination, and conflation with test/retest calculations. Firstly, the time lapse between patient and researcher completion of the ADRe Profile may have triggered some of the issues typically affecting test–retest reliability assessment, such as the carryover effect of the second response being affected by the person’s initial response, or real change in the observed problems over time [102]. This natural fluctuation in clinical signs and symptoms precluded formal (and complex) statistical assessment of test/retest reliability. Secondly, while the raters are assumed to operate independently [66], the second rating in this study occurred in a conversation between the nurse and the patient (Rater 1 and Rater 2).

6. Conclusion

Addressing the ADR problem requires multifaceted approaches, including the development and validation of comprehensive patient-reported outcome decision support questionnaires, alongside the integration of WHO-endorsed practice for managing polypharmacy. By addressing these challenges, healthcare systems can enhance patient safety and optimise the allocation of resources, ultimately mitigating the adverse impacts of ADRs on public health.

The ADRe Profile is suitable for use with older patients in primary care who live at home. Intended users of the instrument had a good understanding of the meaning of ADRe Profile questions and were able to process and record meaningful answers to questions. ADRe Profile responses were sufficiently reliable to be used as a basis for further investigations, prescriber referral and clinical actions.

This paper has described how the validity and reliability of the ADRe Profile were assessed. As part of a wider study, it laid the groundwork for subsequent feasibility testing and a RCT. This approach may serve as a methodological precedent or template for future studies seeking to evaluate the efficacy of novel instruments in healthcare settings.

7. Implications for Nursing Management

To ensure that avoidable problems related to prescribed medicines are not overlooked, managers should facilitate introduction of patient-centred symptom monitoring, with decision support. The safety of older patients prescribed multiple medicines necessitates using a comprehensive instrument, such as the ADRe Profile. Managers should be confident that reliability and validity data are available for instruments adopted.

Data Availability Statement

All relevant data are included in this manuscript or the Supporting Information. Further details are available from the author on request.

Disclosure

The funders had no role in study design, data collection and analysis, interpretation of data, decision to publish or

preparation of the manuscript. The preliminary version of this manuscript was published as a Thesis [103].

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

All authors have equally contributed to the design and writing of the manuscript. Conceptualization: Sue Jordan and Vera Logan; investigation: Vera Logan, Sue Jordan and Neil Carter; writing – original draft: Vera Logan; writing – review and editing: Sue Jordan, David Hughes, Adam Turner and Neil Carter.

Funding

This work was supported by Swansea University (Centenary Research Excellence Scholarship), sponsor reference RIO 034-20.

Acknowledgements

We thank the participants who took part in our study.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information 1. Supporting Information 1: Cognitive interviews—suggested changes to the ADRe Profile and Supporting information.

Supporting Information 2. Supporting Information 2: Cognitive interview participant comments.

Supporting Information 3. Supporting Information 3: Content validity index by ADRe Profile sections.

Supporting Information 4. Supporting Information 4: Between-group differences in problems experienced.

Supporting Information 5. Supporting Information 5: Inter-rater reliability—kappa results per category.

Supporting Information 6. Supporting Information 6: Full results for inter-rater reliability.

References

- [1] I. R. Edwards and J. K. Aronson, "Adverse Drug Reactions: Definitions, Diagnosis, and Management," *Lancet* 356, no. 9237 (2000): 1255–1259, [https://doi.org/10.1016/S0140-6736\(00\)02799-9](https://doi.org/10.1016/S0140-6736(00)02799-9).
- [2] World Health Organisation, "Ageing and Health," (2022), <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health>.
- [3] N. Parameswaran Nair, L. Chalmers, G. M. Peterson, B. J. Bereznicki, R. L. Castelino, and L. R. Bereznicki, "Hospitalization in Older Patients Due to Adverse Drug Reactions-The Need for a Prediction Tool," *Clinical Interventions in Aging* 11 (2016): 497–505, <https://doi.org/10.2147/CIA.S99097>.
- [4] W. N. Insani, C. Whittlesea, H. Alwafi, K. K. C. Man, S. Chapman, and L. Wei, "Prevalence of Adverse Drug Reactions in the Primary Care Setting: A Systematic Review and Meta-Analysis," *PLoS One* 16, no. 5 (2021): e0252161, <https://doi.org/10.1371/journal.pone.0252161>.
- [5] A. Haerdtlein, E. Debold, M. Rottenkolber, et al., "Which Adverse Events and Which Drugs Are Implicated in Drug-Related Hospital Admissions? A Systematic Review and Meta-Analysis," *Journal of Clinical Medicine* 12, no. 4 (2023): 1320, <https://doi.org/10.3390/jcm12041320>.
- [6] J. C. Veeren and M. Weiss, "Trends in Emergency Hospital Admissions in England Due to Adverse Drug Reactions: 2008–2015," *Journal of Pharmaceutical Health Services Research* 8, no. 1 (2017): 5–11, <https://doi.org/10.1111/jphs.12160>.
- [7] E. L. M. Jennings, K. D. Murphy, P. Gallagher, and D. O'Mahony, "In-hospital Adverse Drug Reactions in Older Adults; Prevalence, Presentation and Associated Drugs-A Systematic Review and Meta-Analysis," *Age and Ageing* 49, no. 6 (2020): 948–958, <https://doi.org/10.1093/ageing/afaa188>.
- [8] M. A. Sakiris, M. Sawan, S. N. Hilmer, R. Awadalla, and D. Gnjdjic, "Prevalence of Adverse Drug Events and Adverse Drug Reactions in Hospital Among Older Patients With Dementia: A Systematic Review," *British Journal of Clinical Pharmacology* 87, no. 2 (2021): 375–385, <https://doi.org/10.1111/bcp.14417>.
- [9] R. A. Elliott, E. Camacho, D. Jankovic, M. J. Sculpher, and R. Faria, "Economic Analysis of the Prevalence and Clinical and Economic Burden of Medication Error in England," *BMJ Quality and Safety* 30, no. 2 (2021): 96–105, <https://doi.org/10.1136/bmjqs-2019-010206>.
- [10] E. Braithwaite, O. M. Todd, A. Atkin, et al., "Interventions for Reducing Anticholinergic Medication Burden in Older Adults—A Systematic Review and Meta-Analysis," *Age and Ageing* 52, no. 9 (2023): afad176, <https://doi.org/10.1093/ageing/afad176>.
- [11] J. A. Cole, D. C. Gonçalves-Bradley, M. Alqahtani, et al., "Interventions to Improve the Appropriate Use of Polypharmacy for Older People," *Cochrane Database of Systematic Reviews* 10 (2023): <https://doi.org/10.1002/14651858.CD008165.pub5>.
- [12] M. R. Blum, B. T. G. M. Sallevelt, A. Spinewine, et al., "Optimizing Therapy to Prevent Avoidable Hospital Admissions in Multimorbid Older Adults (OPERAM): Cluster Randomised Controlled Trial," *BMJ* 374 (2021): n1585, <https://doi.org/10.1136/bmj.n1585>.
- [13] R. Holland, C. Bond, D. P. Alldred, et al., "Evaluation of Effectiveness and Safety of Pharmacist Independent Prescribers in Care Homes: Cluster Randomised Controlled Trial," *BMJ* 380 (2023): e071883, <https://doi.org/10.1136/bmj-2022-071883>.
- [14] J. A. Cooper, C. A. Cadogan, S. M. Patterson, et al., "Interventions to Improve the Appropriate Use of Polypharmacy in Older People: A Cochrane Systematic Review," *BMJ Open* 5, no. 12 (2015): e009235, <https://doi.org/10.1136/bmjopen-2015-009235>.
- [15] S. Tecklenborg, C. Byrne, C. Cahir, L. Brown, and K. Bennett, "Interventions to Reduce Adverse Drug Event-Related Outcomes in Older Adults: A Systematic Review and Meta-Analysis," *Drugs & Aging* 37, no. 2 (2020): 91–98, <https://doi.org/10.1007/s40266-019-00738-w>.
- [16] D. S. Atmaja, S. Yulistiani, E. Zairina, and E. Zairina, "Detection Tools for Prediction and Identification of Adverse

- Drug Reactions in Older Patients: A Systematic Review and Meta-Analysis,” *Scientific Reports* 12, no. 1 (2022): 13189, <https://doi.org/10.1038/s41598-022-17410-w>.
- [17] L. B. Mokkink, C. A. C. Prinsen, L. M. Bouter, H. C. W. de Vet, and C. B. Terwee, “The CoConsensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) and How to Select an Outcome Measurement Instrument,” *Brazilian Journal of Physical Therapy* 20, no. 2 (2016): 105–113, <https://doi.org/10.1590/bjpt-rbf.2014.0143>.
 - [18] A. Y. Chen and A. S. Whigham, “Validation of Health Status Instruments,” *ORL: Journal for Oto-Rhino-Laryngology and its Related Specialties* 66, no. 4 (2004): 167–172, <https://doi.org/10.1159/000079873>.
 - [19] J. Khadka, C. Hutchinson, R. Milte, et al., “Assessing Feasibility, Construct Validity, and Reliability of a New Aged Care-Specific Preference-Based Quality of Life Instrument: Evidence From Older Australians in Residential Aged Care,” *Health and Quality of Life Outcomes* 20, no. 1 (2022): 159, <https://doi.org/10.1186/s12955-022-02065-y>.
 - [20] E. Mayo-Wilson, P. Montgomery, S. Hopewell, G. Macdonald, D. Moher, and S. Grant, “Developing a Reporting Guideline for Social and Psychological Intervention Trials,” *British Journal of Psychiatry* 203, no. 4 (2013): 250–254, <https://doi.org/10.1192/bjp.bp.112.123745>.
 - [21] R. Heale and A. Twycross, “Validity and Reliability in Quantitative Studies,” *Evidence-Based Nursing* 18, no. 3 (2015): 66–67, <https://doi.org/10.1136/eb-2015-102129>.
 - [22] C. B. Terwee, C. A. C. Prinsen, M. G. Ricci Garotti, A. Suman, H. C. W. de Vet, and L. B. Mokkink, “The Quality of Systematic Reviews of Health-Related Outcome Measurement Instruments,” *Quality of Life Research* 25, no. 4 (2016): 767–779, <https://doi.org/10.1007/s11136-015-1122-4>.
 - [23] B. B. Frey, *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, 4 (SAGE Publications, Incorporated, 2018).
 - [24] A. J. Stevens, K. R. Abrams, J. Brazier, R. Fitzpatrick, and R. J. Lilford, “Factors That Limit the Number, Progress and Quality of Randomised Controlled Trials: A Systematic Review,” in *The Advanced Handbook of Methods in Evidence Based Healthcare* (SAGE Publications, Limited, 2001), 38, <https://doi.org/10.4135/9781848608344.n3>.
 - [25] S. Jordan, M. Coleman, B. Hardy, and D. Hughes, “Assessing Educational Effectiveness: the Impact of a Specialist Course on the Delivery of Care,” *Journal of Advanced Nursing* 30, no. 4 (1999): 796–807, <https://doi.org/10.1046/j.1365-2648.1999.01176.x>.
 - [26] S. Jordan, C. Tunnicliffe, and A. Sykes, “Minimizing Side-Effects: The Clinical Impact of Nurse-Administered ‘Side-Effect’ Checklists,” *Journal of Advanced Nursing* 37, no. 2 (2002): 155–165, <https://doi.org/10.1046/j.1365-2648.2002.02064.x>.
 - [27] S. Jordan, J. Knight, and D. Pointon, “Monitoring Adverse Drug Reactions: Scales, Profiles, and Checklists,” *International Nursing Review* 51, no. 4 (2004): 208–221, <https://doi.org/10.1111/j.1466-7657.2004.00251.x>.
 - [28] S. Jordan, M. Gabe, L. Newson, et al., “Medication Monitoring for People With Dementia in Care Homes: The Feasibility and Clinical Impact of Nurse-Led Monitoring,” *The Scientific World Journal* 2014 (2014): 843621–843711, <https://doi.org/10.1155/2014/843621>.
 - [29] S. Jordan, M. E. Gabe-Walters, A. Watkins, et al., “Nurse-led Medicines’ Monitoring for Patients With Dementia in Care Homes: A Pragmatic Cohort Stepped Wedge Cluster Randomised Trial,” *PLoS One* 10, no. 10 (2015): e0140203, <https://doi.org/10.1371/journal.pone.0140203>.
 - [30] S. Jordan, T. Banner, M. Gabe-Walters, et al., “Nurse-Led Medicines’ Monitoring in Care Homes Study Protocol: A Process Evaluation of the Impact and Sustainability of the Adverse Drug Reaction (ADRe) Profile for Mental Health Medicines,” *BMJ Open* 8, no. 9 (2018): e023377, <https://doi.org/10.1136/bmjopen-2018-023377>.
 - [31] S. Jordan, H. Prout, N. Carter, et al., “Nobody Ever Questions—Polypharmacy in Care Homes: A Mixed Methods Evaluation of a Multidisciplinary Medicine’s Optimisation Initiative,” *PLoS One* 16, no. 1 (2021): e0244519, <https://doi.org/10.1371/journal.pone.0244519>.
 - [32] M. E. Gabe and S. E. Jordan, “Development and Clinical Gains of Nurse-Led Medication Monitoring Profiles,” *Journal of Nursing Management* 22, no. 3 (2013): 331–349, <https://doi.org/10.1111/jonm.12067>.
 - [33] M. E. Gabe, F. Murphy, G. A. Davies, I. T. Russell, and S. Jordan, “Medication Monitoring in a Nurse-Led Respiratory Outpatient Clinic: Pragmatic Randomised Trial of the West Wales Adverse Drug Reaction Profile,” *PLoS One* 9, no. 5 (2014): e96682, <https://doi.org/10.1371/journal.pone.0096682>.
 - [34] R. Jones, C. Moyle, and S. Jordan, “Nurse-Led Medicines Monitoring: A Study Examining the Effects of the West Wales Adverse Drug Reaction Profile,” *Nursing Standard* 31, no. 14 (2016): 42–53, <https://doi.org/10.7748/ns.2016.e10447>.
 - [35] V. Logan, A. Bamsey, N. Carter, D. Hughes, A. Turner, and S. Jordan, “Clinical Impact of Implementing a Nurse-Led Adverse Drug Reaction Profile in Older Adults Prescribed Multiple Medicines in UK Primary Care: A Study Protocol for a Cluster-Randomised Controlled Trial,” *Pharmacy* 10, no. 3 (2022): 52, <https://doi.org/10.3390/pharmacy10030052>.
 - [36] L. B. Mokkink, M. Boers, C. P. M. van der Vleuten, et al., “COSMIN Risk of Bias Tool to Assess the Quality of Studies on Reliability or Measurement Error of Outcome Measurement Instruments: A Delphi Study,” *BMC Medical Research Methodology* 20, no. 1 (2020): 293, <https://doi.org/10.1186/s12874-020-01179-5>.
 - [37] J. Kottner, L. Audigé, S. Brorson, et al., “Guidelines for Reporting Reliability and Agreement Studies (GRRAS) Were Proposed,” *Journal of Clinical Epidemiology* 64, no. 1 (2011): 96–106, <https://doi.org/10.1016/j.jclinepi.2010.03.002>.
 - [38] P. Campanelli, “Testing Survey Questions,” in *International Handbook of Survey Methodology*, ed. J. Hox, E. de Leeuw, and D. Dillman (Mahwah, NJ: Erlbaum, 2008), 176–200.
 - [39] D. Collins, “Cognitive Interviewing Practice,” *SAGE Research Methods* (2015): <https://methods.sagepub.com/Book/cognitive-interviewing-practice/n1.xml>.
 - [40] A. A. García, “Cognitive Interviews to Test and Refine Questionnaires: Cognitive Interviews,” *Public Health Nursing* 28, no. 5 (2011): 444–450, <https://doi.org/10.1111/j.1525-1446.2010.00938.x>.
 - [41] H. Boeije and G. Willis, “The Cognitive Interviewing Reporting Framework (CIRF): Towards the Harmonization of Cognitive Testing Reports,” *Methodology* 9, no. 3 (2013): 87–95, <https://doi.org/10.1027/1614-2241/a000075>.
 - [42] G. Willis, “Cognitive Interviewing,” *SAGE Research Methods* (2005): <https://methods.sagepub.com/book/cognitive-interviewing/n4.xml>.
 - [43] I. B. Baena, *Cognitive Interviewing Methodology*, ed. K. Miller (Wiley, 2014).
 - [44] D. A. Dillman, “The Logic and Psychology of Constructing Questionnaires,” in *International Handbook of Survey*

- Methodology*, ed. J. Hox, E. de Leeuw, and D. Dillman (Mahwah, NJ: Erlbaum, 2008), 161–175.
- [45] G. B. Willis, *Reducing Survey Error through Research on the Cognitive and Decision Processes in Surveys* (Research Triangle Institute, 1999).
 - [46] G. Willis and J. Lessler, *Question Appraisal System QAS-99* (Research Triangle Institute, 1999), <https://docplayer.net/147787-Question-appraisal-system-qas-99.html>.
 - [47] G. B. Willis, *Analysis of the Cognitive Interview in Questionnaire Design* (Oxford University Press, 2015).
 - [48] J. Boita, A. Bolejko, S. Zackrisson, et al., “Development and Content Validity Evaluation of a Candidate Instrument to Assess Image Quality in Digital Mammography: A Mixed-Method Study,” *European Journal of Radiology* 134 (2021): 109464, <https://doi.org/10.1016/j.ejrad.2020.109464>.
 - [49] P. Bobos, J. C. MacDermid, E. C. Boutsikari, E. A. Lalone, L. Ferreira, and R. Grewal, “Evaluation of the Content Validity Index of the Australian/Canadian Osteoarthritis Hand Index, the Patient-Rated Wrist/Hand Evaluation, and the Thumb Disability Exam in People With Hand Arthritis,” *Health and Quality of Life Outcomes* 18, no. 1 (2020): 302, <https://doi.org/10.1186/s12955-020-01556-0>.
 - [50] X. Chen, L. Luo, L. Jiang, et al., “Development of the Nurse’s Communication Ability With Angry Patients Scale and Evaluation of Its Psychometric Properties,” *Journal of Advanced Nursing* 77, no. 6 (2021): 2700–2708, <https://doi.org/10.1111/jan.14788>.
 - [51] M. R. Lynn, “Determination and Quantification of Content Validity,” *Nursing Research* 35, no. 6 (1986): 382–386, <https://doi.org/10.1097/00006199-198611000-00017>.
 - [52] C. F. Waltz and R. B. Bausell, *Nursing Research: Design, Statistics, and Computer Analysis* (F.A. Davis Co, 1981).
 - [53] D. F. Polit and C. T. Beck, *Nursing Research: Generating and Assessing Evidence for Nursing Practice*, 10th ed. (Wolters Kluwer, 2017).
 - [54] C. A. Wynd, B. Schmidt, and M. A. Schaefer, “Two Quantitative Approaches for Estimating Content Validity,” *Western Journal of Nursing Research* 25, no. 5 (2003): 508–518, <https://doi.org/10.1177/0193945903252998>.
 - [55] J. R. Landis and G. Koch, “The Measurement of Observer Agreement for Categorical Data,” *Biometrics* 33, no. 1 (1977): 159–174, <https://doi.org/10.2307/2529310>.
 - [56] K. M. L’Ecuyer, D. S. Subramaniam, and C. Reangsing, “Development of the Preceptor Self-Assessment Tool and Use of the Content Validity Index,” *The Journal of Continuing Education in Nursing* 51, no. 10 (2020): 469–476, <https://doi.org/10.3928/00220124-20200914-09>.
 - [57] S. K. Grove and J. Gray, *Understanding Nursing Research: Building an Evidence-Based Practice*, 8th ed. (Elsevier, 2023), <https://evolve.elsevier.com/cs/product/9780323826419?role=student>.
 - [58] A. R. Newman and P. R. Helft, “Reliability and Validity of a Tool to Assess Oncology Nurses’ Experiences With Prognosis-Related Communication,” *Oncology Nursing Forum* 42, no. 1 (2015): 64–73, <https://doi.org/10.1188/15.ONF.64-73>.
 - [59] E. E. Ruel, W. E. Wagner, and B. J. Gillespie, *The Practice of Survey Research: Theory and Applications* (SAGE Publications, Inc, 2016).
 - [60] C. A. Gershon, A. N. Yagapen, A. Lin, D. Yanez, B. C. Sun, and D. B. Diercks, “Inter-Rater Reliability of the HEART Score,” *Academic Emergency Medicine* 26, no. 5 (2019): 552–555, <https://doi.org/10.1111/acem.13665>.
 - [61] D. A. Grimes and K. F. Schulz, “Uses and Abuses of Screening Tests,” *Lancet* 359, no. 9309 (2002): 881–884, [https://doi.org/10.1016/S0140-6736\(02\)07948-5](https://doi.org/10.1016/S0140-6736(02)07948-5).
 - [62] A. Swift, R. Heale, and A. Twycross, “What Are Sensitivity and Specificity?” *Evidence-Based Nursing* 23, no. 1 (2020): 2–4, <https://doi.org/10.1136/ebnurs-2019-103225>.
 - [63] G. Shan, “Sample Size Calculation for Agreement Between Two Raters With Binary Endpoints Using Exact Tests,” *Statistical Methods in Medical Research* 27, no. 7 (2018): 2132–2141, <https://doi.org/10.1177/0962280216676854>.
 - [64] H. Hong, Y. Choi, S. Hahn, S. K. Park, and B. J. Park, “Nomogram for Sample Size Calculation on a Straightforward Basis for the Kappa Statistic,” *Annals of Epidemiology* 24, no. 9 (2014): 673–680, <https://doi.org/10.1016/j.annepidem.2014.06.097>.
 - [65] L. A. Goodman and W. H. Kruskal, “Measures of Association for Cross Classifications,” *Journal of the American Statistical Association* 49, no. 268 (1954): 732–764, <https://doi.org/10.1080/01621459.1954.10501231>.
 - [66] J. Cohen, “A Coefficient of Agreement for Nominal Scales,” *Educational and Psychological Measurement* 20 (1960): 37–46, <https://doi.org/10.1177/001316446002000104>.
 - [67] Datapharm, <https://www.datapharm.com/resource-hub/>.
 - [68] T. Pringsheim, D. Gardner, D. Addington, et al., “The Assessment and Treatment of Antipsychotic-Induced Akathisia,” *Canadian Journal of Psychiatry* 63, no. 11 (2018): 719–729, <https://doi.org/10.1177/0706743718760288>.
 - [69] D. F. Polit, C. T. Beck, and S. V. Owen, “Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations,” *Research in Nursing & Health* 30, no. 4 (2007): 459–467, <https://doi.org/10.1002/nur.20199>.
 - [70] British National Formulary, *British Medical Association and the Royal Pharmaceutical Society of Great Britain*, (2023).
 - [71] C. A. Lantz and E. Nebenzahl, “Behavior and Interpretation of the κ Statistic: Resolution of the Two Paradoxes,” *Journal of Clinical Epidemiology* 49, no. 4 (1996): 431–434, [https://doi.org/10.1016/0895-4356\(95\)00571-4](https://doi.org/10.1016/0895-4356(95)00571-4).
 - [72] Swansea University, (2024), <https://www.swansea.ac.uk/adre/>.
 - [73] M. E. Echevarría-Guanilo, N. Gonçalves, and P. J. Romanoski, “Psychometric Properties of Measurement Instruments: Conceptual Basis and Evaluation Methods-Part II,” *Texto & Contexto Enfermagem* 28 (2019): <https://doi.org/10.1590/1980-265X-TCE-2017-0311>.
 - [74] S. Stephens-Davidowitz, *Everybody Lies: Big Data, and What They Tell Us About Who We Really Are* (Dey Street Books, 2017).
 - [75] L. A. Clark and D. Watson, “Constructing Validity: New Developments in Creating Objective Measuring Instruments,” *Psychological Assessment* 31, no. 12 (2019): 1412–1427, <https://doi.org/10.1037/pas0000626>.
 - [76] N. W. Hudson, I. Anusic, R. E. Lucas, and M. B. Donnellan, “Comparing the Reliability and Validity of Global Self-Report Measures of Subjective Well-Being with Experiential Day Reconstruction Measures,” *Assessment* 27, no. 1 (2020): 102–116, <https://doi.org/10.1177/1073191117744660>.
 - [77] N. Gao, M. S. Rahaman, W. Shao, and F. D. Salim, “Investigating the Reliability of Self-Report Data in the Wild: The Quest for Ground Truth” (2021), <https://doi.org/10.48550/arxiv.2107.00389>.

- [78] C. H. Dushey, L. J. Bornstein, M. M. Alexiades, and G. H. Westrich, "Short-Term Coagulation Complications Following Total Knee Arthroplasty," *The Journal of Arthroplasty* 26, no. 8 (2011): 1338–1342, <https://doi.org/10.1016/j.arth.2010.11.007>.
- [79] S. Alazzawi, N. V. Bardakos, S. G. Hadfield, U. Butt, Z. H. Beer, and R. E. Field, "Patient-reported Complications After Elective Joint Replacement Surgery: Are They Correct?" *Journal of Bone and Joint Surgery British Volume* 94, no. 8 (2012): 1120–1125, <https://doi.org/10.1302/0301-620X.94B8.29040>.
- [80] J. N. Greenbaum, L. J. Bornstein, S. Lyman, M. M. Alexiades, and G. H. Westrich, "The Validity of Self-Report as a Technique for Measuring Short-Term Complications after Total Hip Arthroplasty in a Joint Replacement Registry," *The Journal of Arthroplasty* 27, no. 7 (2012): 1310–1315, <https://doi.org/10.1016/j.arth.2011.10.031>.
- [81] Electronic Medicines Compendium (Emc) a, "Risperidone," (2011), <https://www.medicines.org.uk/emc/product/8207/smpc>.
- [82] Electronic Medicines Compendium b, "Gabapentin," (2015), <https://www.medicines.org.uk/emc/product/2362/smpc>.
- [83] D. Hughes, M. Jordan, P. A. Logan, et al., "Looking for the 'Little Things': A Multi-Disciplinary Approach to Medicines Monitoring for Older People Using the Adre Resource," *Geriatrics* 5, no. 4 (2020): 79–13, <https://doi.org/10.3390/geriatrics5040079>.
- [84] World Health Organisation, *Medication without Harm: Policy Brief* (World Health Organization, 2023), <https://iris.who.int/bitstream/handle/10665/376212/9789240062764-eng.pdf?sequence=1>.
- [85] World Health Organisation, "Medication Safety in Polypharmacy," (2019), <https://www.who.int/docs/default-source/patient-safety/who-uhc-sds-2019-11-eng.pdf>.
- [86] D. Stacey, F. Légaré, K. Lewis, et al., "Decision Aids for People Facing Health Treatment or Screening Decisions," *Cochrane Database of Systematic Reviews* 2017, no. 4 (2017): CD001431, <https://doi.org/10.1002/14651858.CD001431.pub5>.
- [87] J.-P. Michel, M. Leonardi, M. Martin, and M. Prina, "WHO's Report for the Decade of Healthy Ageing 2021–30 Sets the Stage for Globally Comparable Data on Healthy Ageing," *The Lancet. Healthy Longevity* 2, no. 3 (2021): e121–e122, [https://doi.org/10.1016/S2666-7568\(21\)00002-7](https://doi.org/10.1016/S2666-7568(21)00002-7).
- [88] K. Jones, H. Weatherly, S. Birch, et al., "Unit Costs of Health and Social Care 2022," (2022), [https://kar.kent.ac.uk/100519/1/Unit%2520Costs%2520of%2520health%2520and%2520Social%2520Care%25202022%2520\(amended%252013%2520uly%25202023\).pdf](https://kar.kent.ac.uk/100519/1/Unit%2520Costs%2520of%2520health%2520and%2520Social%2520Care%25202022%2520(amended%252013%2520uly%25202023).pdf).
- [89] E. R. Burns, J. A. Stevens, and R. Lee, "The Direct Costs of Fatal and Non-Fatal Falls Among Older Adults—United States," *Journal of Safety Research* 58 (2016): 99–103, <https://doi.org/10.1016/j.jsr.2016.05.001>.
- [90] J. A. Campbell, A. Venn, A. Neil, M. Hensher, M. Sharman, and A. J. Palmer, "Diverse Approaches to the Health Economic Evaluation of Bariatric Surgery: A Comprehensive Systematic Review," *Obesity Reviews* 17, no. 9 (2016): 850–894, <https://doi.org/10.1111/obr.12424>.
- [91] L. Hazell and S. A. W. Shakir, "Under-Reporting of Adverse Drug Reactions: A Systematic Review," *Drug Safety* 29, no. 5 (2006): 385–396, <https://doi.org/10.2165/00002018-200629050-00003>.
- [92] P. Shuttleworth, J. Baker, and E. Clark, "Under-Reporting of Gastrointestinal Bleeding Associated With Anticoagulant Use Using the UK Yellow Card Scheme," *International Journal of Clinical Pharmacy* 45, no. 4 (2023): 1014–1018, <https://doi.org/10.1007/s11096-023-01601-0>.
- [93] J. B. Schroll, E. Maund, and P. C. Gotzsche, "Challenges in Coding Adverse Events in Clinical Trials: A Systematic Review," *PLoS One* 7, no. 7 (2012): e41174, <https://doi.org/10.1371/journal.pone.0041174>.
- [94] D. W. Light and J. R. Lexchin, "Pharmaceuticals as a Market for 'Lemons': Theory and Practice," *Social Science & Medicine* 268 (2021): 113368, <https://doi.org/10.1016/j.socscimed.2020.113368>.
- [95] L. Waddell and M. Taylor, "A New Self-Rating Scale for Detecting Atypical or Second-Generation Antipsychotic Side Effects," *Journal of Psychopharmacology* 22, no. 3 (2008): 238–243, <https://doi.org/10.1177/0269881107087976>.
- [96] T. Wykes, J. Evans, C. Paton, et al., "What Side Effects Are Problematic for Patients Prescribed Antipsychotic Medication? the Maudsley Side Effects (MSE) Measure for Antipsychotic Medication," *Psychological Medicine* 47, no. 13 (2017): 2369–2378, <https://doi.org/10.1017/S0033291717000903>.
- [97] T. W. A. Schoenmakers, M. Teichert, M. Wensing, and P. A. G. M. de Smet, "Evaluation of Potentially Drug-Related Patient-Reported Common Symptoms Assessed during Clinical Medication Reviews: A Cross-Sectional Observational Study," *Drug Safety* 40, no. 5 (2017): 419–430, <https://doi.org/10.1007/s40264-017-0504-7>.
- [98] R. Lim, L. K. Ellett, E. E. Roughead, P. Y. Cheah, and N. Masnoon, "Patient-Reported Questionnaires to Identify Adverse Drug Reactions: A Systematic Review," *International Journal of Environmental Research and Public Health* 18, no. 22 (2021): 11877, <https://doi.org/10.3390/ijerph182211877>.
- [99] M. McCaffery, K. Miller, and I. B. Baena, "Nursing the Patient in Pain. 1983," *Cognitive Interviewing Methodology*, 2nd ed. (Wiley, 2014).
- [100] A. Sayer, *Method in Social Science* (Routledge, 1992).
- [101] R. Tourangeau, "The Survey Response Process From a Cognitive Viewpoint," *Quality Assurance in Education* 26, no. 2 (2018): 169–181, <https://doi.org/10.1108/QAE-06-2017-0034>.
- [102] D. F. Polit, "Getting Serious about Test–Retest Reliability: A Critique of Retest Research and Some Recommendations," *Quality of Life Research* 23, no. 6 (2014): 1713–1720, <https://doi.org/10.1007/s11136-014-0632-9>.
- [103] V. Logan, "Medication Monitoring in General Practice: Clinical Impact of Implementing a Nurse-Led Adverse Drug Reaction (ADRe) Profile in Older Adults With 5 or More Prescribed Medicines," *Swansea, Wales, UK* (2024): <https://cronfa.swan.ac.uk/Record/cronfa66314>, <https://doi.org/10.23889/SUthesis.66314>.